UNIVERSITÀ DEGLI STUDI DI PALERMO

Dottorato di Ricerca in Tecnologie e Scienze per la Salute dell'Uomo
Dipartimento di Promozione della Salute, Materno-Infantile, di Medicina Interna e Specialistica
d'Eccellenza "G.D'Alessandro"
SSD: BIO/18

DEVELOPMENT OF NONTOXIC BIO-ADHESIVES FOR WET ENVIRONMENTS

IL DOTTORE
FRANCESCA VENTURELLA

IL COORDINATORE
PROF. MAURIZIO LEONE

IL TUTOR
PROF. SALVATORE FEO

CO-TUTOR
DOTT.SSA CATERINA ALFANO

CICLO XXXIII
ANNO CONSEGUIMENTO TITOLO 2021

The research work here presented was funded by Fondazione Ri.MED through Grant **"CheMISt"- Computational Molecular Design e Screening -** Patto per il Sud Regione Siciliana (CUP G77B17000110001),

and Grant "**Potenziamento Infrastruttura di Ricerca "GMP Facility, Laboratori di Ricerca e Servizi Diagnostici e Terapeutici IRCCS-ISMETT**" - PO FESR Sicilia 2014/2020. Azione 1.5.1. (CUP G76G17000130007) - Partnership IRCCS-ISMETT and Fondazione Ri.MED.

The research activities were carried out in collaboration with Dr Rosa Passantino of Istituto di Biofisica, Consiglio Nazionale delle Ricerche, Palermo, and Prof Annalisa Pastore of Department of Basic and Clinical Neuroscience, King's College London, London.

Acknowledgments

A recently published study, in the Washington Post, has highlighted the fact that being grateful in life is not just a matter of education, of bon ton. Gratitude is one of the most precious and beneficial emotions we can feel and give to others.

There are many people to whom, today, I want to say thank you to, because each of them, in their own way, has contributed in giving life, light and form to these three years of work, permitting me to grow, not only professionally but, also, in a personal way.

I want to express my gratitude to my two supervisors, Prof. Salvatore Feo and Dr. Caterina Alfano, without whom this thesis work would not have been possible; their precious advice, their example, and the time they've dedicated to me, has increased in me the love and the dedication towards research.

I thank the Coordinator of my PhD, Prof. Maurizio Leone, for the availability, patience and kindness with which he has always responded to the requests of us PhD students.

A warm thanks goes to Prof. Annalisa Pastore and her research group, for the affection and hospitality shown during the months of study spent in London.

I sincerely thank the Structural and Biophysical Biology group of Dr. Caterina Alfano and, in particular, Raffaele Sabbatella, Elisa Monaca and Maria Agnese Morando, for the affection, support and friendship shown during these three years.

Finally, my thanks go to my friends and to my whole family for always supporting and sustaining me in the most difficult moments and rejoicing with me in the most beautiful ones.

Table of Contents

1. Introduction

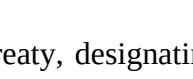
Every second, in our body, millions of new cells are born to replace those lost during cellular aging or after a tissue damage caused by traumas and disease. This process occurs naturally and harmonically in our body and exploits the regenerative capacities of the human body to permit it to repair itself and to restore its normal function. Each gear of the machinery works in syntony with the others to make the machine efficient in its totality. For this reason, the human body is often identified as a perfect machine and has often been raised up to divine beauty by Renaissance-art geniuses.

However, it can encounter an advanced pathological deterioration in an irreversible way, making physiological tissue regeneration impractical. For this reason, it has been necessary to applicate and blend together the principles of medicine, physics, engineering, chemistry, and biology in a new science: tissue engineering[1,2]. This uses a brand-new approach which is able to revolutionize the entire medicine: the development of biomaterials on which cells can attach and proliferate, reconstructing a damaged tissue or an entire organ.

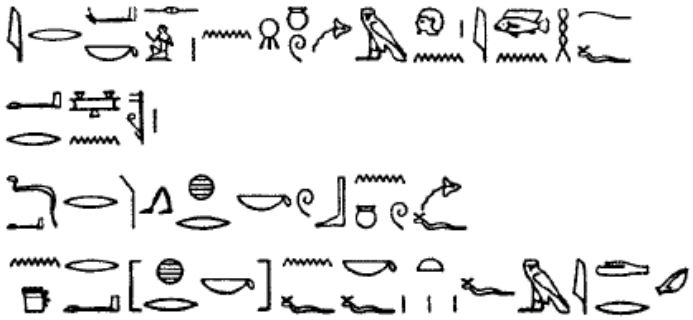
At present, though, this innovative approach is still at the beginning, and organ and tissue transplant is still the most employed practice to promptly cure organ failures and to recover vital functions that may guarantee patient's survival.

In this context, the great challenge is, to offer patients a better after-surgery process, which is often complicated by tissue distortions due to suture used in conventional surgery. In the last decades, many efforts are therefore dedicated to the development of bioadhesives, able to compensate the great limitations given by conventional surgery. Indeed, bioadhesives are not invasive and could guarantee a better tissue integration as well as to quicken post-surgical recovery.

History teaches us that our ancestors had already clear the limits of wound closure by suture, and rectified it by resorting to adhesive bandages to guarantee a better tissue adhesion. The first known reference of surgical stitching in history comes to us from ancient Egypt. The Smith's Papyrus[3] (Fig. 1.1), dating back to the XVII century, a.C, is surely one of the most interesting historical findings and it represents the first proof of surgical practice known to us. The first treaty of Egyptian medical practice is divided in 48 cases of severe and chronicle wounds with an exhaustive and unusual description of the wounds and of their medicament. Inside the treaty each case presents an examination

followed by a diagnosis and treatment of injuries. The new word, that we find often repeated in the treaty, designating the noun stitching is 

. In the case the wound would not close perfectly, after having practised suture, the ancient Egyptian surgeon would recur to an adhesive plaster, applying grease, honey, lint and rigid linen bandages over the wound. Even though ancient Egyptians did not know the complex chemical mechanism that rules adhesion between two surfaces chemically different, they had the merit to have conceived the first adhesive bandages obtained from Egyptian territory, *provider of fodder and producer of many drugs* (**Homerus, Odyssey, book IV, 219-232**).



"If you examine a man having a wound in the top of his eyebrow, penetrating to the bone, you should palpate his wound and draw together for him his gash with stitching"

Figure 1.1 *Case X*, Edwin Smith's Papyrus XVII century a.C in hieroglyphic script

Grateful to this past knowledge for the contribution given to today's surgery approach, modern medicine has brought man to endeavour always more competitive reconstructive processes, which are part of non-invasive surgical techniques. In that sense bioadhesives, offer a strong and valid contribute in tissue engineering and regeneration, facilitating tissue integration, reducing mechanic irritation and inflammation, and improving tissue regeneration. One of the most representative example is that of using bioadhesives in the treatment of comminuted bone fractures. The complexity in restoring the bone structure from such severe fragmentation, could indeed be overcome by using bioadhesives as scaffold to restore bone architecture. Unfortunately, as already said, the use of

biomaterials, in surgical field, is a branch still young and continues to encounter many obstacles, the most representative of which is water. Indeed, the chemical-physical features of H_2O, the most abundant and vital molecule in our organism (70%), make very challenging the development of a valid bioadhesive, able to make tissues adhere in an efficient and durable way in wet environments. In the last decades, several haemostatic agents and synthetic adhesives have been developed, but their use is limited to only few surgical procedures. As a result, sutures, staples, grafts, *etc.*, are still mostly used in surgical field to close wounds, despite the side effects due to their use, such as high infection rate, loss of pus and exudates, low motility, no bleeding control and tissue damage. Haemostatic agents are also widely used, in particular to promptly manage bleeding events, although they are known to meet only few needs.

1.1 Haemostatic agents

During a surgical operation, the most important thing is certainly dictated by the time of action. Many are the complications that can arise in the operating room and to intervene promptly becomes vital. The haemostatic agents are among the first materials chosen by surgeons to rapidly stop haemorrhage during a surgical operation. From the Greek αίμοστατικός, αἷμα «blood» e στατικός «able to stop», the haemostatics are agents with a mechanic action or chemical mechanism that accelerate the coagulation process of the blood. Among the haemostatics, to undertake a mechanic action, we find porcine gelatine, bovine collagen, oxided regenerated cellulose and polysaccharide spheres[4]. Protein-derived haemostatic agents, such as thrombin, fibrin-glue and MeTRO hydrogels, are instead representative examples of haemostatics with a chemical mechanism.

Porcine gelatine: Available in the form of sponges or powder, the porcine gelatine are produced starting from the collagen extracted from skin, bone and connective tissue of pigs[5]. One of the problems connected to this kind of haemostatics is the risk of swelling with probable damage of the structures close to the application site. From a clinical point of view, the risk of infections is potentially high, causing toxic shock syndrome by blocking drainage, foreign body reactions, and tendon fixation by prolonged fibrosis[6]. Another critical aspect in the use of these materials regards Muslim religion: the word *haram*, from the arabic "*forbidden*", prohibits Muslims to eat pork meat, which is

considered impure, and so also the application, for biomedical use, of pig-derived materials.

Bovine collagen: The use of bovine collagen as haemostatic overcomes the religion issue, but still share most of the safety warnings seen in porcine gelatine. This haemostatic is made by the fibrous protein main component of connective tissue, bones, tendons, cartilage, and it is commercially available under the form of sheets and flours. Available different techniques have been adapted to make the use of these materials more safety and to avoid the risk of swelling[7], but their use remains not indicated in closed cavities and in neurosurgical field.

Oxidized regenerated cellulose: This kind of haemostatics is derived from cotton cellulose. It creates an acidic environment and is bactericidal. Applied to the haemorrhagic affected area, oxidized regenerated cellulose is absorbed rapidly changing into gel, adapting itself perfectly to the geometric irregularities of the wound. Despite the acknowledged efficacy, this haemostatic presents most of the drawbacks mentioned before, comprising also the risk of foreign body reactions[6].

Polysaccharide spheres: These are another example of haemostatic agents with mechanic action of vegetable origin. Polysaccharide spheres are able to stop and control haemorrhage by simply absorbing blood and local liquids, increasing their mass and so the risks associated with it, such as swelling, with resulting pressure on tissues and surrounding organs. Due to their polysaccharide composition, particular attention is to be given to limit their dosage in diabetic patients, so to not incur into an increase of the load of glucose in the blood and into glycosuria.

Thrombin: This enzyme actively participates to the fibrin clots formation. The clot formation is given essentially by the transformation of the fibrinogen protein, soluble, in fibrin, insoluble protein that detains in its interior the corpuscular elements of the blood. The thrombin enzyme accelerates this fibrinogen to fibrin conversion. Commercially available as bovine, human plasma and recombinant, the thrombin can give rise to adverse reactions due to the transmission of viral diseases and be cause of allergies for patients even in its recombinant form due to hamster and snake proteins used in manufacturing[6].

Fibrin glue: Ready-to-use fibrin is available in form of glue. These haemostat and sealant glues exhibit a better performance on dry surfaces than in wet tissues, limiting their use in many clinical applications.

MeTRO: Methacryloyl-substituted Tropoelastin (MeTRO) is a human protein-based surgical sealant recently developed as an elastic hydrogel-based sealant, and it seems to promise high hopes for its application in neurosurgical field and in the treatment of chronical wound. In in- vitro and in- vivo studies on rats, MeTRO have demonstrated an efficient biodegradability and a wide antibacterial action spectrum[8,9].

1.2 Synthetic adhesives

In the last two decades many synthetic adhesives have been developed. They are known to have been able to solve many critical surgical procedures, even though they need to be improved for their non-negligible secondary effects[6]. Among these, are the cyanoacrylate adhesives, also known as *Superglue*, which were discovered by H. Coover, in 1999[10]. They have been widely used in different surgical fields, as the thoracic, vascular, plastic, maxillofacial, orthopaedic and in gynaecological ed urogenital fields[11]. The monomeric general structure of cyanoacrylates, exposed to humidity, polymerizes rapidly at room temperature and forms long polymeric chains, that are able to bind surfaces together. The polymerization mechanism of cyanoacrylates (Fig. 1.2) is anionic and the chain initiation takes place by attack of a nucleophilic species (hydroxide ions) to ethylene carbon.

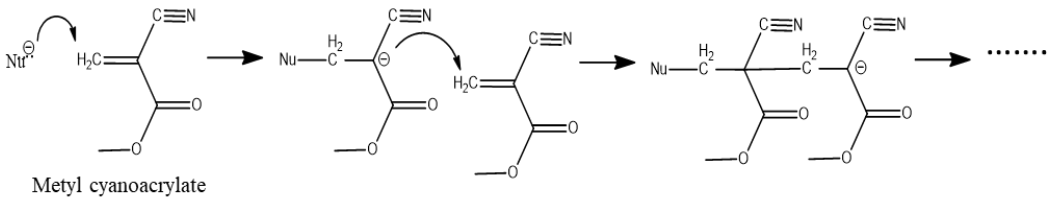


Figure 1.2. Polymerization mechanism of methyl cyanoacrylate

Cyanoacrylate adhesives have been widely used until 2004[12] because of their easy preparation, and recognized high tissue adhesion features. Today, they have a more limited employment to a topical use because they are associated with an inadequate bonding strength and a poor biocompatibility due to potentially dangerous degradation products formed, during the exothermic polymerization reaction[6]. This brings to classify

cyanoacrylates-based products as external synthetic adhesives[13]. Another example of synthetic adhesives is represented by Polyurethanes, synthetic polymers, produced from isocyanate and organic compounds containing multiple hydroxyl groups (polyalcohols) in the presence of catalyst (Fig. 1.3).

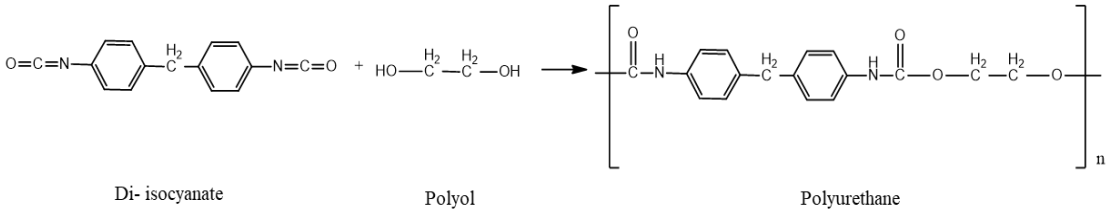


Figure 1.3. Synthesis of a typical Polyurethane

These synthetic polymers have been used in the most varied medical applications as bladders, catheters, cardiovascular applications, wound dressing and pacemakers[14]. Although they are prepared starting from an *ad hoc* and controlled chemical process, their biocompatibility and efficiency in biomedical applications are still lacking. Later, in fact, it has been discovered that polyurethanes are associated with different disadvantages like high incidence of infection, poor tissue adherence, slow resorption, pseudoarthrosis, tissue necrosis and delay in scar formation[15]. Less side-effects have been observed for polyethylene glycol (PEG), synthetic polymer widely employed both as drug delivery system in pharmaceutical formulation, and as scaffold for regenerative medicine[16].

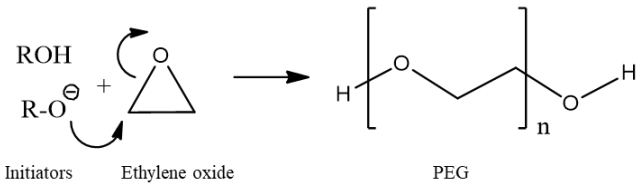


Figure 1.4. Polymerization of ethylene oxide

PEG is an inert amphiphilic polymer prepared from ethylene oxide polymerization (Fig.1.4). Mammals do not have enzymes able to metabolize PEG's polymers[17], so no side products are produced following their application. Their employment however has to be avoided in close spaces that could lead to nerve compression[6,16], since the tendency of these polymers to cause swelling. Another feature that surely should not be underestimated is the *clearance* by the renal and hepatic pathway. Precautions should therefore be taken for those patients with kidney and liver disfunction.

Despite the potential benefits of some tissue adhesives, commercially available products have limited applications due to limited functionality[18]. Adhesives found on the market are developed to meet few needs and none of these can be used for all potential applications. The future bioadhesives have to satisfy specifical criteria, including the ability to interact with surfaces with different make-up, maintain adhesive properties, possess antimicrobial behaviour, resist to mechanical strains, be biocompatible that is, it should not only be toxic but it should integrate with the surrounding tissue and work in wet environment. To this purpose, scientists have focused on natural adhesive proteins and many efforts are invested in developing tissue adhesives inspired by Nature, which has always been a main source of inspiration for human-being. The *Sunu*, ancient Egyptian word, indicating, "*he who cures whom suffers*", had a limited anatomical knowledge and the only medical school he attended, at the time, was Nature. Man began to cure himself using spontaneous herbs and observing animals who did that instinctively, and still today Nature can represent a great source of inspiration to develop products that can improve human-being life.

1.3 Adhesion in wet environments observed in Nature

Nature is full of animals for which adhesion is crucial for their life as a part of their defensive strategy or for attachment on different surfaces. Many examples are found in land animals such as spiders, flies and geckos but, above all, in animals of the sea world, like sea cucumbers, sea stars, barnacles, marine tubeworms and mussels[19].

The different life styles that characterize each marine organism, like also the biology of the species, influence strongly the chemical composition and the properties of the secreted adhesive. In particular, there are two types of adhesion: i) permanent adhesion involves the secretion of an adhesive that hardens over time and it is characteristic of sessile organisms that remain in the same place for all their life, such as rock attachments of barnacles and mussels; ii) temporary adhesion allows a greater exploration and a simultaneous adhesion of different surfaces, as happens for sea stars[20,21]. In both cases, adhesives are secreted in seawater environment with high pH and ionic strength, thus already proving that, in principle, they could be used in tissue adhesion. Among all sea organisms able to secrete adhesive products, mussels have captured more the attention of many researchers because of the unique property of their secreted adhesive proteins able

to toughly and durably attach to marine surfaces also in very adverse tidal conditions. Thus many efforts have been done to create strong materials suitable to act as biomedical adhesives[22], starting from studying the composition of mussel adhesive proteins, and trying to mimic their adhesive properties[23–25].

1.3.1 Mussel Adhesive Proteins

Mussel is a very surprising complete organism[26–28] (fully anatomy description is reported in ANNEX I). Its adhesive proteins are located in the byssus, the secreted filaments that mussels use to attach to surfaces (Fig.1.5). A single thread of a mussel's byssus, is composed of three parts (core, cuticle and plaque), and it is secreted by three different types of cells, situated in the mussel foot: collagen gland, phenol gland and enzyme gland. The byssus core consists of fibrous biopolymer, rich in collagen proteins, known as *PreCol*, covered by a thin and strong protective cuticle with a granular structure. The byssus ends with a proteins rich plaque, containing the adhesive mussel foot proteins (*mfps*)[29,30]. These adhesive proteins are like a glue, so strong that it permits mussels to adhere firmly to rocks even in extreme sea condition.

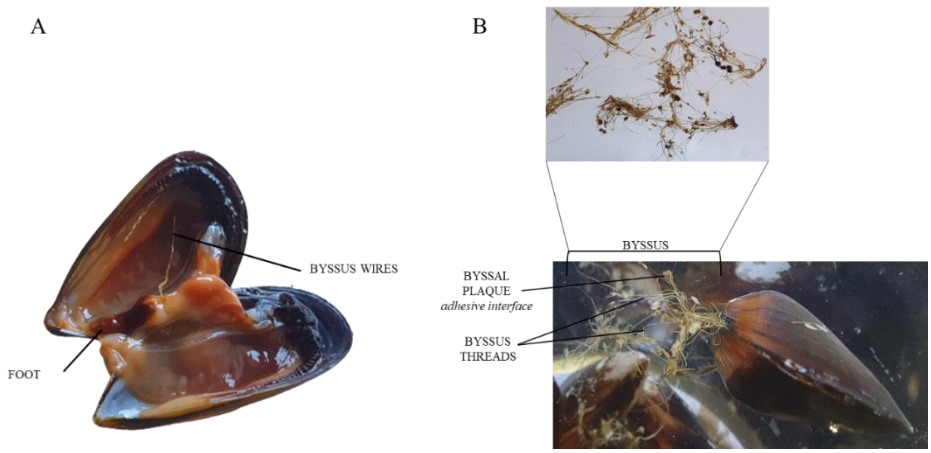


Figure 1.5: A. Mussel foot and byssus wires. **B.** Mussel byssus threads

Although not yet fully characterized, the byssus has as many as twenty different known protein components, designated as (*mfp*)-2, to -6. A great part of the *mfps* contain in their sequence a big quantity of cathecolic amino acid, the 3,4-Dihydroxyphenylalanine (Dopa) that is produced by post-translational modification of tyrosine residues and contribute to adhesion[31–34]. The proteins *mfp-3* and *mfp-5* are the richest in Dopa and they can be found

principally in the interface between the adhesive plaque and the substrate[20]. Dopa explicates its adhesive properties in its reduced form by: exposing the two hydroxyl groups to the anchoring surface, forming hydrogen bonds and coordination with metal ions (metal- chelating), forming bidentate complexes with the surface, π-interactions, π-π stacking and hydrophobic interactions[35–42]. A reduced form of Dopa is maintained thanks to the presence of cysteine residues in *mfps* sequence, which are able to keep reduced Dopa by disulfide bridges formation[43]. Anyhow, Dopa oxidized form, Dopa-quinone, is essential to guarantee the strong attachment to the surface since it allows the plaque solidification by cross-linking. To maintain both reduced and oxidized forms of Dopa, mussels impose a specific chemical environment on the distal depression of its foot[29] (Fig. 1.6). Chemicals are released into the cavity formed by the foot distal depression to maintain an acidic environment and redox regulation. It follows the release of the adhesive proteins that undergo to a liquid-liquid phase separation (LLPS)[44]. When the mussel foot is lift from surface, cross-linking interactions prevail because adhesive proteins equilibrate themselves with sea water, which is oxidizing and with a more basic pH value. In these conditions, the plaque is assembled followed by the formation of a cuticle over the plaque that undergoes to solidify[45].

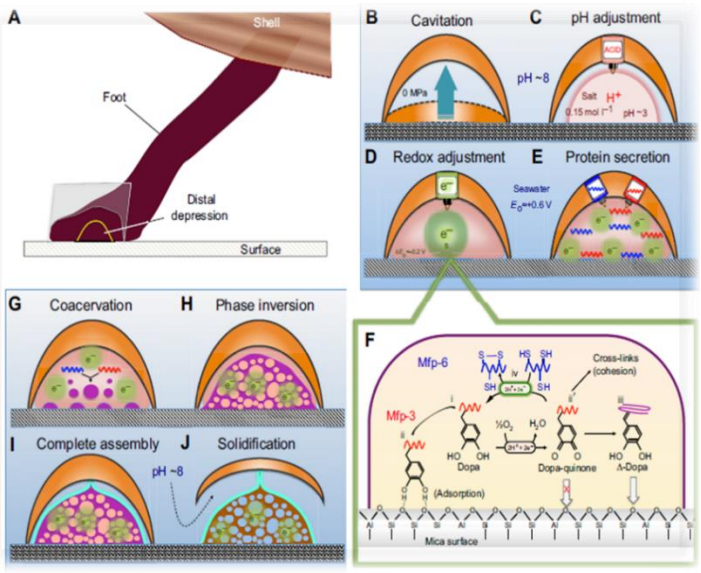


Figure 1.6. Plaque protein deposition by the foot. Image taken from J. H. Waite, *J. of Experimental Biology.* 2017, **220**, 517-530[29]. **A.** The inverted cup shape represents a 2D cross-sectional view of the distal depression in contact with a target surface for adhesion. Labelled panels show the stages of plaque protein deposition. **B.** Cavitation or creation of negative pressure between foot and surface. **C.** Secretion of acid to pH as low as ~2. **D.** Redox regulation. **E.** Release of adhesive proteins (red and blue) and adsorption to target surface. **F.** Redox activity driven by the difference between the high pH and O_2 concentration of seawater versus the low pH and abundance of electron donors in the plaque: (i) Adhesive protein *Mfp-3* (red) or *Mfp-5* is deposited and (ii) binds to the target surface; as Fe^{3+} or O_2 leaks in, some Dopa is oxidized to (ii′) Dopa-quinone, which is poor in adhesion but active in cross-linking (cohesion). (iii) At increasing pH, Dopa-quinone can self-reduce to Δ-Dopa (with conformational effects). (iv) Dopa can also be rescued by reduction using electrons from thiolates in Mfp-6. Not shown is the favorable scavenging reduction of O_2 by *Mfp-6* (blue) thiolates. **G.** Coacervation: proteins undergo fluid–fluid phase separation. **H.** Coacervate/water phase inversion. **I.** Plaque assembly is completed and a protective cuticle is added over the plaque. **J.** Solidification of fluid.

Starting from the strong and concrete idea that the adhesive key, in the mussel foot proteins, resides in Dopa, many researchers have developed dopaminated polymers to mimic those produced by mussels[46–50]. In this regard, an example is that of Dopa-modified poly-(ethyleneglycol)- (PEG) developed by Messersmith group[51]. Molecular weight analysis, performed on monosubstituted PEG-Dopa, revealed that up to six Dopa molecules react with one another to form a network junction, resulting in a rapid hydrogel formation. More recently, Zeng *et al.* have produced an adhesive hydrogel by UV irradiation of a mix composed by three polymers: 3,-4-dihydroxyphyenylanine-acrylamide (DOPA-AA), poly-(ethylene-glycol)-diacrylate (PEGDAA), and thiolated-chitosan (CSS)[52]. The hydrogel showed an improvement of the adhesive properties mainly due to the interactions occurring between the catechol group of DOPA-AA and the thiol group of CSS. However, cytotoxicity tests performed on L929 cells showed reduced cells viability, probably induced by the catechol oxidation.

Only in recent years, researches start to investigate on the role of other possible factors that, together with Dopa, would contribute to keep unique the adhesive properties of *mfps*. One of the possible factors is that demonstrated by Ding Haibing group: a possible implication of lipids in mussel adhesion[53]. The presence of fat materials in the adhesive proteins of many organisms, such as geckos, flies and sea species, is not a novelty, but the molecular mechanisms behind this is still unclear. Fatty acids, found in the byssus thread and in the mussel's plaque, could improve adhesion to the surface, wiping away, for example, water molecules, offering electrostatic attractions, and also acting as a protective barrier on the just formed plaque just, against bacteria. The role of the other amino acids found inside *mfps* are also under investigation. In fact, it is not to be overlooked the presence of Lysine, close to Dopa, whose mechanism is not clear yet, but it seems to mediate many of the interactions with *mfps*-metals, like Fe^{3+} ion, and it seems to be involved in coacervated formation[23].

More studies are needed to fulfill the lack of a complete knowledge in the chemical and biological complex mechanism that presides over the adhesion in *mfps.* In particular, there are not structural information on any *mfps*, probably due to difficulties in producing sufficient quantity of proteins for structural studies. The only structural studies, found today in literature, concern mussel foot proteins from *Perna viridis*[54,55] and rely only on homology modelling approach. A deeper structural characterization would definitively

help to better understand the complex redox mechanism around Dopa and the effective role of the other amino acid residues involved in the adhesion to surface.

1.3.2 *Perna viridis* foot proteins

The only structural studies on *mfps* conducted so far regard adhesive foot proteins from *Perna viridis*[54–61], a green mussel that lives in the south east of the Asian coast and in the Indian Ocean. The adhesive proteins of these green mussels are named after *Perna viridis foot proteins* (*Pvfps*). *Perna viridis* foot contains at least eight kinds of potential precursor proteins, designated as (Pvfp)-a to -h. Among the precursors, Pvfp-g and Pvfp-h are found at relatively high levels in the phenol gland. These proteins are considered to be essential components in the adhesive plaque, which mediates adhesion of the thread to the substrate surface according to Ohkawa group[62]. That which differentiates *Perna viridis* from *Mytilus edulis* species is: the colour of the shell, green or bluish green, the absence of the anterior adductor muscle[63,64], genome organization[65], and much more reduced information in literature about *Pvfps* adhesion mechanisms[66]. On the other hand, more structural information are available for *Pvfps*[54,55].

Petrone *et. al*[54] reported the structural modelling of three *Pvfps*: *Pvfp-3α*, *Pvfp-6* and *Pvfp-5β*. These proteins are time regulated, so *Pvfp-5β* is first secreted,[67] followed by *Pvfp-3α* and *Pvfp-6*. Each type of *Pvfps* has a unique function and contributes differently to adhesion on surface. According to Petrone's model studies, the *Pvfp-5β* has a more elongated shape with a high percentage of β-strands in their secondary structure, connected by loops, and also of a high content of Dopa. The *Pvfp-3α* is the smallest of the three and it has a globular shape, a high percentage of cysteine residues and β-strands, and also an α-helical region. The *Pvfp-6* adopts an extended, non-globular conformation and has a good percentage of cysteine and proline residues, a low percentage of β-strands, a small α-helix and a high percentage of loops. The *Pvfp-5β* is surely, among the three proteins, the most interesting because it is the first to be secreted by mussels and to initiate an interaction with sea surface. Recently researchers have investigated more carefully the role of positively charged residues -close to Dopa and Tyrosine in *Pvfp-5β*[68]. As suggested from their Molecular dynamics (MD) simulations, all of the studied configurations share a prolate ellipsoidal shape with Dopa and positively charged basic residues staying on the protein surface. Confirmed by MD simulations and Nuclear

Magnetic Resonance (NMR) spectroscopy there is also a spatial correlation between positively charged residues, like Lysine and Arginine. Especially Lysine, the most abundant positively charged residues in *Pvfp-5β*, seems to be related with an improving of peptide adhesion through electrostatic attraction with the surface, facilitating the interpeptide cross-linking rather than hydrogen bond formed by Dopa with surface. Even though Lysine seems to be enhance *Pvfp-5β* adhesion with surface, other researcher group have investigated the role of Lysine on cohesion[69] and how mussels control the balance between surface adhesion and cohesion. Unlike surface adhesion, Lysine seems to impede cross-linking interaction between Dopa molecules and Fe^{3+}, known to be crucial in mussel foot proteins solidification, due to the electrostatic repulsion that occurs between Lysine and Fe^{3+}. Adhesive properties of Tyr-*Pvfp-5β* and Dopa-*Pvfp-5β* was also tested and compared by Bilotto *et. al*[70]. Their findings support the idea that the presence of Dopa molecules is not sufficient condition to generate adhesion with surface as emerging from various studies on mussel foot proteins. In fact, Tyr-*Pvfp-5β* and Dopa-*Pvfp-5β* resulted to possess comparable adhesion behaviour ability.

1.4 Thesis Aims

As outlined in the preceding sections, the structural characterization of *Mussel foot proteins* could overcome today's lack of a really complete understanding of how these adhesive proteins work and interact with wet surfaces. The understanding of adhesive protein folding can strongly contribute in designing biomaterials behaving the same as the natural source to be employed in different areas of human health. In particular, the aim is to structurally characterize the recently identified byssal plaque adhesive protein from the *Perna viridis* mussel species, Pvfp-5β; the first to be secreted by mussels and to initiate an interaction with sea surface. This could give important adhesion behaviour information, crucial for new bioadhesives design. In this regard, recombinant Pvfp-5β was produced to jump the difficulty in recovering *Pvfps* proteins from their native source, and, more important, to deal with a no Dopa form of Pvfp-5β and explore adhesive features other than Dopa. It followed a deep analysis of the protein structure by Nuclear Magnetic Resonance (NMR), Mass Spectrometry (MS) and Circular Dichroism (CD) techniques, and an extensive probing of its cell adhesion properties.

2. Materials and Methods

2.1 Recombinant proteins production

The *Perna Viridis* foot protein 5β (Pvfp-5β) used in this research work, was produced through DNA recombinant technique, which is nowadays the best way to achieve high amount of high-quality proteins for structural studies[71]. The pioneers of this efficient technique were Stanley Cohen and Herbert Boye, who, in 1973, inserted the gene for frog ribosomal RNA into a plasmid sequence, and then transferred the modified plasmid in *E.coli* cells. The cells thus transformed presented the frog ribosomal RNA, proving for the first time that genetic material from one organism can be artificially introduced into the genome of another organism and then replicated and expressed by that other organism[72]. The new technique immediately captured the interest of pharmaceutical industries, and it soon became the most used technique for the production of proteins for test analyses in sufficient quantity and at lowest cost. The approach was so successful that, in 1982, recombinant human insulin was released to the market as treatment for diabetes[73], and, nowadays, recombinant proteins are routinely used in biomedical and biotechnological research, and more than 130 recombinant proteins are approved for medical use[74]. The production of recombinant proteins required several steps and expertise. The first step is to isolate the DNA sequence encoding for the protein of interest, known as complementary DNA (cDNA). This is allowed through Polymerase Chain Reaction technique (PCR) which is enable to amplify many copies of cDNA[75,76]. Afterwards follows a process known as *ligation,* in which, the cDNA fragment, is inserted into an expression vector, a plasmid, which allows its replication. Once the complete vector is obtained, it must be incorporated into a host cell. The most widely used host organism is *Escherichia coli*, identified by Theodor Escherich in 1885, but eukaryotic systems can also be used. The choice is crucial and it has to be made to ensure the proper folding of the protein of interest, and to sustain it in the intact and functional state[77]. Transformed host cells are then selected thanks to a selected marker carried by the vector, generally an antibiotic resistance gene, that allows survival in medium containing antibiotic only to those cells that internalized the vector. Furthermore, a vector has to include an inducible promoter activable by a chemical inducer, the origin of the replication and the correct translation initiation sequence. In this way, the vector is able

to co-opt the translation machinery of the host cell and express the protein encoded by the inserted cDNA. To recover and purify the recombinant protein, the cDNA is often engineered to introduce a nucleotide sequence encoding for a tag or a fused protein upstream or downstream the sequence encoding for the protein of interest. Generally a polyhistidine tag is used, and it permits to purify the protein by Immobilized Metal Affinity Chromatography (IMAC)[78], where the metal ions of the chromatographic resin are strongly coordinated by histidine residues of the fused tag (Fig. 2.1). The target protein is then collected by passing Imidazole through the column since Imidazole competes with the His-tag for binding to the Metal-charged resin.

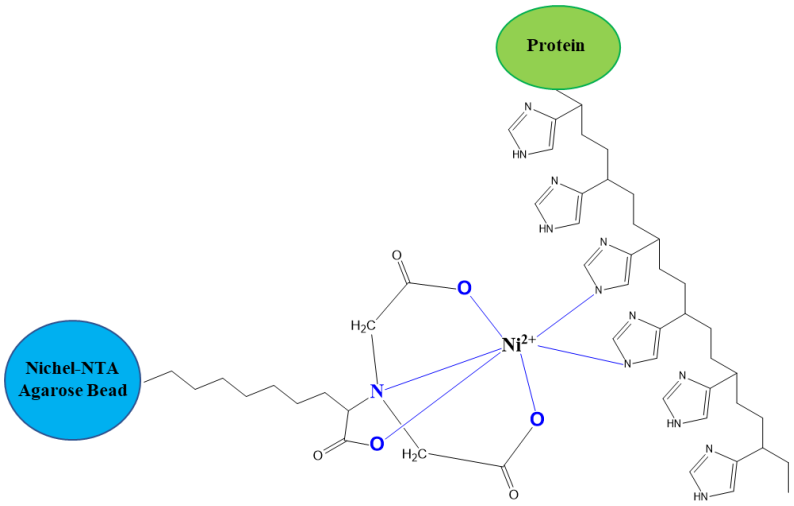


Figure 2.1. Affinity purification of polyhistidine-tagged recombinant proteins by IMAC

2.1.1 Preparation of Pvfp-5β expression vectors

Two different bacterial expression vectors were prepared for Pvfp-5β (Fig. 2.2). The plasmid hereafter referred as phTt-Pvfp-5β (sequence details in ANNEX II.1), encoded for full length Pvfp-5β fused at its N-terminal with an exa-histidine tag, followed by Thioredoxin and the target cleavage site for Tobacco Etch Virus (TEV) protease. The Pvfp-5β cDNA was synthesised by Integrated DNA Technologies Inc., and it was inserted into an in-house derivative of pMAL-c5x (New England Biolabs Inc.) by Sequence Ligation Independent Cloning (SLIC) method[79].

The phTt plasmid was amplified by PCR using KAPA-Hifi Hotstart Polymerase (Kapa Biosystems, Roche), digested with 20 U of DpnI (New England Biolabs Inc.) at 37°C for 2 h to cleave the methylated DNA template, and purified with Monarch® DNA Cleanup

kit (New England Biolabs Inc.). The linear phTt and the synthetized insert were then mixed in a 1:4 ratio in NEB2.1 buffer (New England Biolabs Inc), and the mix was treated with 0.5 U of T4 DNA polymerase for 1 min at room temperature to allow the exonuclease activity of the enzyme. Annealing was performed by incubation of the mix in ice for 15 minutes, and the mixture was then inserted into DH5α competent cells by standard heat shock protocol.

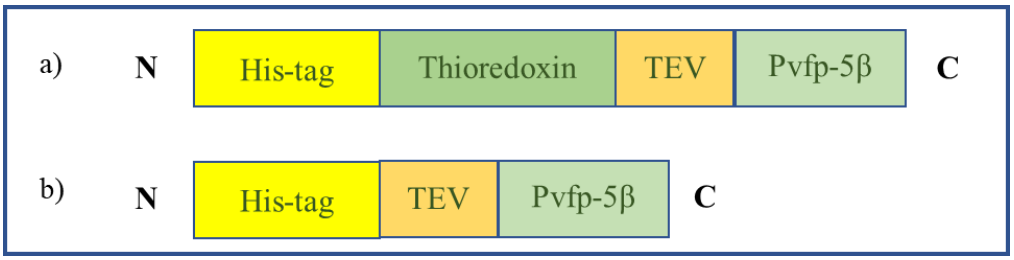


Figure 2.2: Domains organization of Pvfp-5β constructs: **a)** phTt-Pvfp-5β construct; **b)** pHT-Pvfp-5β construct

A second plasmid, hereafter referred as pHT-Pvfp-5β, was obtained via PCR deletion of the thioredoxin gene of phTt-Pvfp-5β, resulting in a cDNA encoding for full length Pvfp-5β fused at its N-terminal with an exa-histidine tag, followed by TEV protease cleavage site (Fig. 2.2). The template plasmid was amplified using a pair of mutually complementary primers annealing to the DNA regions flanking the thioredoxin coding sequence (Fig. 2.3). The mixture composition for 50 µl of PCR reaction is reported in Table 2.1, while thermo-cycling conditions are reported in Table 2.2.

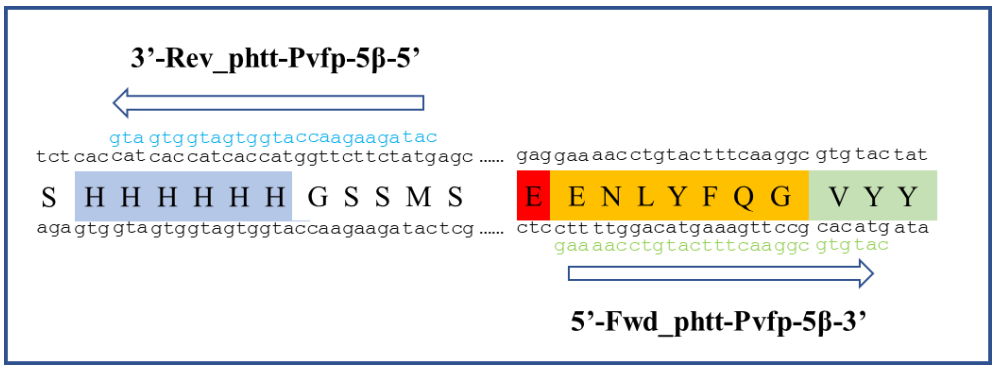


Figure 2.3: Primers design to delete Thioredoxin site from phtt-Pvfp-5β construct. Highlighted in green the sequence for forward primer, and in cyan the sequence for reverse primer

Component	Volume (µl)
MilliQ H_2O	18.5
10X PCR reaction buffer (QuikChange Lightning Buffer, Agilent Technologies)	5
QuikSolution reagent (Agilent Technologies)	1.5
dNTPs mix	1
12.5ng/µl Forward primer	10
12.5ng/µl Reverse primer	10
~10ng/µl Template DNA (phTt-Pvfb5beta)	4
Taq DNA Polymerase (QuikChange Lightning Enzyme, Agilent Technologies)	1

Table 2.1: Mixture composition of 50µl PCR reaction performed to delete thioredoxin encoding region from phTt-Pvfp-5β plasmid.

Step	Temperature (°C)	Time	No. of Cycle
Initial denaturation	95	2min	1
Denaturation	95	20s	18
Annealing	53	10s	
Extension	68	150s	
Final extension	68	5min	1

Table 2.2: PCR thermal-cycling conditions for deletion of thioredoxin encoding region from phTt-Pvfp-5β plasmid.

The resulting nicked DNA vector was digested by adding 2 µl of DpnI to the PCR mixture and incubating it at 37°C for 5 minutes. 0.8% (w/v) Agarose gel electrophoresis was run at 90V for 30 minutes to confirm the product and to separate it from unspecific PCR products. The linearized new plasmid was extracted from the gel and purified using the Gel Extraction Kit from Sigma-Aldrich. Purified product was then phosphorylated by adding 7.5µl 2X Rapid Ligation Buffer (Promega), 1µL of 10U/ml T4 PnK1 (Promega) and 1.5µl MilliQ water to a final volume of 15 µL, and incubating it at 37 °C for 1 hour. DNA Ligase (Promega) was then used to circularize the linear mutated DNA. 1µl of 3U/µl T4 DNA Ligase (Promega), 2.5µl of 2X Ligase Buffer, and 1.5µl of MilliQ water were added to DNA fragments and the mixture was first incubated at room temperature for 2 hours and then used to transform XL10-Gold Ultracompetent Cells by standard heat shock transformation protocol. The nucleotide sequence of the plasmid was confirmed by Sanger sequencing (BMR Genomics).

2.1.2 Protein expression tests

In order to achieve the better Pvfp-5β protein expression, different *E.coli* cell strains and conditions, as induction temperature and time, were tested. For this purpose, were chosen: i) BL21(DE3)pLysS cells (Agilent Technologies) that provide a tighter control of protein expression avoiding basal expression of the protein before induction. pLysS strains, in fact, express T7 lysozyme, which further suppresses basal expression of T7 RNA polymerase prior to induction, thus stabilizing pET recombinants encoding target proteins that affect cell growth and viability; ii) ArcticExpress(DE3) cells (Agilent Technologies) that address the common bacterial gene expression hurdle of protein insolubility. ArcticExpress competent cells have been engineered to co-express the cold-adapted chaperonins Cpn10 and Cpn60 from the psychrophilic bacterium, *Oleispira antarctica*. These chaperonins improve the processing at lower temperatures of the recombinant protein to be expressed, potentially increasing the yield of its active and soluble fraction; and iii) Origami-2(DE3)pLysS cells (Novagen) that both avoid basal expression of the protein and facilitate disulfide bond formation, thanks to the presence of mutations in glutathione reductase (gor) and thioredoxin reductase (trxB). All three cell strains contain the lambda DE3 prophage that carries the gene for T7 RNA polymerase under control of a lacUV5 promoter, allowing expression of the T7 RNA polymerase to be induced with Isopropil-β-D-1-tiogalattopiranoside (IPTG). For each strain, expression tests were conducted as following. Cells were transformed with pHT-Pvfp-5b plasmid, and then cultured at 37°C in Luria-Bertani (LB) medium complemented with 50 μg/ml ampicillin and with 20 μg/ml gentamycin (Sigma-Aldrich) for Arctic cells. Protein expression was induced by adding 1 mM IPTG following the manufacturer's protocol. Two different expression conditions were tested: 3 hours at 37°C, and overnight at 18°C. The expression conditions that resulted in the higher protein expression level, were then used to produce the protein in large scale. Protein expression tests were verified by SDS-PAGE analysis (NuPAGE™ 4 to 12%, Bis-Tris, Mini Protein Gel, Life Technologies).

2.1.3 Expression of unlabelled protein in large scale

Once the best protein expression conditions have been investigated, by expression tests, Pvfp-5β protein was expressed in large scale. Transformed BL21(DE3)pLysS/Pvfp-5β

cells were inoculated into 5 mL of LB medium, supplemented with ampicillin, and cultured for 8 hours under shaking at 37°C. The pre-start culture was then diluted into 100 mL of fresh LB/ampicillin medium, and incubated under shaking at 37°C until it reached 0.6-0.8 OD_{600nm}. Expression was induced by 1 mM IPTG for 3 hours at 37°C and 250 rpm. After 3 hours of protein expression, the cells were harvested by centrifugation for 30 minutes at 7500 rpm. The cell pellet was stored at -80°C until further processing.

2.1.4 Expression of isotopically labelled protein in large scale

^{15}N,^{13}C labelling is necessary for protein structure determination by Nuclear Magnetic Resonance (NMR), which experiments depend on the correlation of ^{15}N and ^{13}C nuclei. Transformed E. coli BL21(DE3)pLysS/Pvfp-5β cells were grown in 5 mL LB/ampicillin medium for 8 hours at 37°C under shaking. The pre-start culture was then spanned down and resuspended in 100 mL of ^{15}N,^{13}C-enriched M9 minimal medium (^{15}N,^{13}C-M9 MM, Table 2.3) supplemented with ampicillin. The resulted culture was incubated at 37°C overnight, under shaking to allow cells to adapt to the minimal medium. Overnight culture was diluted in 2 liters of fresh ^{15}N,^{13}C-M9 MM supplemented with ampicillin, and cultured at 37°C until it reached 0.6-0.8 $OD_{600\ nm}$. Expression was induced by adding 1 mM IPTG for 3 hours at 37°C and 250 rpm. After 3 hours of protein expression, the cells were harvested by centrifugation for 30 min at 7500 rpm in a centrifuge. The cell pellet was stored at -80°C until further processing.

Component	**Quantity**	**Final concentration**
10X M9 salt solution pH 7.2 (Na_2HPO_4*$2H_2O$ 0.422 M, KH_2PO_4 0.220 M, NaCl 85.55 mM)	100 mL	1X
^{15}N-NH_4Cl	0.7 g	0.7 g/l
^{13}C-Glucose	2.0 g	2 g/l
1 M $MgSO_4$	2 mL	2 mM
1M $CaCl_2$	0.1 mL	0.1 mM
1 mg/mL Biotin	1 mL	1 mg/l
1 mg/mL Thiamine	1 mL	1 mg/l
100X Trace elements stock (Fe^{3+} 3.0 mM, Zn^{2+}0.4 mM, Cu^{2+}0.07 mM, Co^{2+}0.04 mM, H_3BO_3 0.16 mM, Mn^{2+}0.13 μM)	10 mL	1X

Table 2.3. Protocol for 1L of ^{15}N,^{13}C-enriched M9 minimal medium

2.1.5 Protein purification

In all tested expression conditions, Pvfp-5β resulted to be aggregated into cell inclusion bodies. It was then necessary to optimize an extraction and refolding procedure.

The cell pellet was washed twice with 100 mM Tris-HCl, 10 mM EDTA, 5 mM $CaCl_2$ (pH 7.4), and then with 20 mM sodium phosphate buffer at pH 7.4. The washed pellet was resuspended in pre-chilled lysis buffer (20 mM sodium phosphate buffer at pH 7.4, 500 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$, cOmplete EDTA free protease inhibitor-Roche, 10 μg/ml DNase, and 0.5 mg/ml lysozyme). Cells were disrupted by ultrasonic homogenizers (Bandelin HD 2070), and incubated 30 min at 4°C. The inclusion bodies containing recombinant His-tagged Pvfp-5β (HT-Pvfp-5β) were then washed twice in 20 mM sodium phosphate buffer at pH 7.4, 500 mM NaCl, 2 mM DTT, cOmplete EDTA free protease inhibitor-Roche, and 1% Triton X-100. The washed inclusion bodies were solubilized for 2 h at room temperature then overnight at 4°C in 8 M urea, 1 M NaCl, 2 mM DTT, 20 mM sodium phosphate buffer at pH 7.4. The supernatant, containing denatured and reduced protein, was collected by centrifugation at 20,000 g, 4 °C for 30 min, filtered using 0.45 μm filter (Sartorius Stedim Biotech), and loaded onto a 5 ml HisTrap FF crude column prepacked with Ni-Sepharose (GE Healthcare Life Sciences) pre-equilibrated in the same buffer of the protein sample. HT-Pvfp-5β was eluted under denaturing and reducing conditions with a linear gradient from 0 to 500 mM of imidazole in 10 CV at room temperature and was verified by SDS-PAGE analysis (NuPAGE™ 4 to 12%, Bis-Tris, Mini Protein Gel, Life Technologies).

Refolding of the eluted protein was then performed by sequential extensive dialysis at 4°C, firstly in 20 mM sodium phosphate at pH 7.4, 2 M urea, 250 mM NaCl, 2 mM reduced glutathione (GSH) and 0.5 mM oxidised glutathione (GSSG), then in the same buffer with no urea. A final dialysis step was performed in 20 mM sodium phosphate at pH 7.4 and 250 mM NaCl. Protein purity was verified by SDS-PAGE analysis, while protein concentration was assessed by UV spectrophotometric determination at 280 nm (extinction coefficient [ε] = 27570 $M^{-1}cm^{-1}$).

His-tag removal was performed by adding TEV protease in 1:50 ratio to the protein solution, and incubating at room temperature for two hours under shaking. Protein cleavage was verified by SDS-PAGE analysis. Reverse IMAC chromatography was then

conducted to recover cleaved Pvfp-5β. The Protein/TEV mixture was centrifuged at 4700 rpm for 10 minutes at 4°C, filtered using 0.45 μm filter (Sartorius Stedim Biotech), and loaded onto a 5 ml HisTrap FF crude column prepacked with Ni-Sepharose (GE Healthcare Life Sciences), pre-equilibrated in 20 mM sodium phosphate pH 7.4, 250 mM NaCl and 10 mM imidazole. Because of its strong adhesive features, cleaved Pvfp-5β was not collected in the column flow through as expected, but it was eluted with a step gradient from 0 to 100 % of imidazole in 25 CV at room temperature. Extensive dialysis was then performed at 4°C in 20 mM sodium phosphate pH 7.4, 250 mM NaCl to remove imidazole. A last dialysis step was performed in 5% acetic acid, which allowed an easy lyophilization of the protein sample. Protein purity was verified by SDS-PAGE analysis, and protein concentration was assessed by UV spectrophotometric determination at 280 nm (extinction coefficient [ε] = 26080 $M^{-1}cm^{-1}$).

2.2 Cell adhesion, cell spreading, and cytotoxicity tests

2.2.1 Surface coating of Pvfp-5β

Tissue culture treated (hydrophilic surface, Corning), and untreated (hydrophobic surface, Greiner Bio-One International) polystyrene 96-well plates (TCT-PS and TCUT-PS respectively), and tissue culture treated SensoPlate Plus glass bottom 96-well plates (TCT-G, Greiner Bio-One International), were coated with HT-Pvfp-5β or Cell-Tak (BD Bioscience), a natural extract of *mfps*. The amount of coating material used was 7 $\mu g/cm^2$ of well area. Cell-Tak coated wells and uncoated wells were used as positive and negative controls, respectively. Coating with HT-Pvfp-5β and Cell-Tak were performed based on the Cell-Tak manufacturer's instruction (adsorption method). Cell-Tak or HT-Pvfp-5β dissolved in 5% acetic acid was diluted with Milli-Q water up to 19 μM, and 10 μl were placed into each well, and mixed with three-fold volumes of neutral buffer solution (0.1 M sodium bicarbonate, pH 8.3). After incubation at room temperature for 10-16 h, the solution was aspirated. Wells were washed thoroughly with sterile Milli-Q water for 1 h. The coating of attached proteins was visualized by Coomassie-blue staining. Images were taken of two independently coated wells per sample.

2.2.2 Mammalian cell culture

Human HeLa (ACC-673–DSMZ) and murine NIH 3T3 (SIGMA-ALDRICH) cell lines were cultured at 37°C and 5% CO_2 humidified air in Dulbecco's Modified Eagle Medium (DMEM)-high glucose, supplemented with 10% fetal bovine serum (FBS) and 10% bovine calf serum (BCS), respectively, 1% penicillin and streptomycin (10,000 U/mL and 10,000 µg/mL, respectively). Cells were treated with trypsin (SIGMA-ALDRICH) at 80% confluency using the standard protocol, and resuspended into new culture flasks. Cells were passaged every third day.

2.2.3 Cell proliferation, adhesion and spreading on Pvfp-5β coated surface

Cell vitality and proliferation were examined by CellTiter 96® AQueous One Solution Cell Proliferation Assay-(MTS) (PROMEGA). NIH-3T3 and HeLa cells were diluted in serum-containing medium, and seeded into uncoated or HT-Pvfp-5β coated wells of TCT-PS 96-well plates at a density of 5x103/well. After 24 h, 48 h and 72 h, 20 µl of CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega) were pipetted into each well. Plates were incubated at 37°C for 3-4 h in a humidified 5% CO2 atmosphere. Absorbance was read at 490 nm using a multi-well plate reader (BioRadiMarkTMMicroplate Reader). The obtained absorbance values were directly proportional to the number of viable cells in culture. MTS assays were also applied for quantitative cell binding measurement after cell adhesion on coated substrates. HeLa and NIH-3T3 cells were diluted in serum-free medium, and seeded into uncoated and coated wells of TCT-PS, TCUT-PS and TCT-G 96-well plates described before, at a density of 5x104/well. After incubation at 37 °C for 2 h, the unattached cells were aspirated gently with PBS, 200 µl of fresh serum-containing medium were added and MTS assay was immediately performed as described in the cell proliferation assay. Each MTS assay was performed at least in triplicate. In cell spreading assays, HeLa and NIH-3T3 cells were diluted in serum-free medium, seeded at a density of 1.1 × 105/ well into uncoated, and Poly-L-Lysine (PLL) or Pvfp-5β coated wells of an 8-well Nunc Lab-Tek II chamber slide. After incubation at 37°C for 2 h, the unattached cells were aspirated gently with PBS. The adhesion and spreading of cells were analyzed by actin filaments staining by using the green-fluorescent Alexa Fluor™ 488 Phalloidin (Invitrogen). Briefly, the

adherent cells were fixed with 4% paraformaldehyde for 20 min, permeabilized with 0.5% Triton X-100 in PBS for 10 min and blocked for 30 min with 5% NGS (Sigma-Aldrich), 0.1% Triton X-100 in PBS. Alexa Fluor™ 488 Phalloidin (1:250) in blocking solution was added and incubated for 1 h in the dark. Then, cells were rinsed twice with PBS and stained with Hoechst 33342 fluorescent DNA-binding dye at 0.01 mg/mL at room temperature for 10 min. After rinsing three times with PBS, the chamber was removed and the slide sealed for the microscopic inspection. The nuclear morphology and cytoskeleton were observed on a Nikon Eclipse 80i microscope equipped for epifluorescence and recorded by a digital camera system.

2.3 Structural characterization

It is well recognized that the function of any protein is closely related to its structure[80]. To solve the three-dimensional structure of a protein is thus paramount, and allows the understanding at atomic level of the molecular mechanism that drives the biological function of the protein itself. Techniques such as Mass Spectrometry (MS), Dynamic Light Scattering (DLS), and Circular Dichroism (CD), are combined to characterize and describe physical and chemical protein properties, such as estimation of their molecular weight, their particle size and their distribution in solution, their stability, and the assessment of disulfide bridges and post-translational modifications. Nuclear Magnetic Resonance (NMR), X-ray crystallography, and Cryogenic Electron Microscopy (Cryo-EM), are used instead to determine the protein structure, and any binding mode in which it is involved. NMR remains the unique technique that allows to measure protein dynamics in solution.

In the here presented research work, MS, CD and NMR were extensively used to structural characterize Pvfp-5β. MS has become the method of choice for protein identification and quantification, especially when facing with a mixture of proteins. The information that can be achieved by MS are: protein molecular weight, post-translational modification, assessment of disulfide bridges, protein elemental composition, thermochemical values, that is, for example ionization energies and appearance energies, of neutral molecules, ions, and radicals if present. In this project, MS was used to confirm the molecular weight of recombinant Pvfp-5β, and to determine cysteine pairs of disulfide

bridges. CD was used to investigate the secondary structure content in Pvfp-5β in different buffer conditions. Indeed, CD is the first elite technique that can provide protein secondary structure information and can allow to observe conformational changes under the effect of different factors, such as temperature and pH. NMR was chosen as main technique to determine Pvfp-5β three-dimensional structure at atomic resolution, and it will be extensively discussed in the following paragraphs. NMR also allowed a deep investigation of Pvfp-5β dynamics in solution.

2.3.1 Mass spectrometry analysis of HT-Pvfp-5β

Mass spectrometry[81], in contrast to the other spectroscopic techniques, is a destructive method of analysis and thus it is not used to study the interaction between radiation and matter, such as either absorption, emission or scattering of radiation by a molecule. The molecule, under test, is bombarded with a powerful electron beam, giving rise to its ionization in gas phase. The so formed fragmented ions, are unique for each molecule, and are separated according to their mass/charge ratio. ElectroSpray Ionization (ESI) and Matrix-Assisted Laser Desorption Ionization (MALDI) are the most widely employed methods for proteins ionization. The first one allows to work with charged macromolecules in solution, from which an aerosol is produced. Indeed, electrospray ionization process leads to a low molecule fragmentation, resulting particularly suitable for studying large macromolecules like proteins and nucleic acids. Instead MALDI is a ionization technique that involves a support matrix[82], where the molecule of interest is dissolved, and is hit by a laser beam. Since molecule is complexed with matrix, it is released as clasterized form ions, which are analized. This lead to create ions from large molecules with minimal fragmentation.

In this research work, to assess the presence of disulfide bridges, proteolytic digestion approach, coupled with mass spectrometry, was used in HT-Pvfp-5β. Generally, the protein sulfhydryls, not involved in disulfide bridges, are reacted with iodoacetamide, and analyzed by ESI-MS. Follow the proteolytic digestion on modified protein, which is performed by enzymes, such as trypsin and chymotrypsin. These enzymes catalyze the hydrolysis of the peptide bonds, at certain and known cutting sites[83], and led to shorts lengths of amino acid, which are analyzed by Liquid Chromatography coupled with

tandem Mass Spectrometry (LC-MS/MS). The resulting spectra, of protein fragmentation, can be analyzed to identify which cysteines are involved in disulfide bridges.

MS combined with proteolytic digestion was here used to identify the disulfide bridges of refolded HT-Pvfp-5β. The freeze-dried protein (10 nmol) was dissolved in 0.1 mM HCl (pH 4), diluted in 100 mM ammonium bicarbonate (final pH 7.5) and incubated with a 100-fold molar excess of iodoacetamide for 30 minutes in the dark to achieve covalent modification of the cysteine residues not involved in disulfide bonds. The reaction was stopped by removing the unreacted iodoacetamide by SPE cartridge (Chromabond HRX-Macherey-Nagel). An aliquot (0.5 nmol) of the modified HT-Pvfp-5β underwent LC-MS analysis on a Q-TOF premier instrument coupled with an Alliance 695 HPLC pump (Waters) to determinate the number of modified cysteines and verify the homogeneity of the reacted protein. The protein was loaded on an Aeris C4 (150x 2.1 mm) column (Phenomenex) and eluted performing a gradient from 20% to 80% of acetonitrile over 5 minutes. Mass spectra were acquired in positive ion mode over the 1500–3000 m/z range. The modified protein was then lyophilised, dissolved in 100 mM ammonium bicarbonate (pH 7.5) and incubated with proteomic-grade trypsin (enzyme:protein ratio 1:50 w/w) at 37 °C. After 6 hours of reaction, chymotrypsin (enzyme:protein ratio 1:100 w/w) was added to the mixture and the digestion proceeded for further 2 hours and stopped by lowering the pH to 2 with 1 M HCl. Resulting fragments were analysed by the LC-MS system. Peptide separation was achieved with a Aeris C18 (150x 2.1 mm) column (Phenomenex) and a gradient from 5% to 65% of acetonitrile over 30 minutes. Mass spectra were in positive ion mode over the 500–2000 m/z range; most abundant species underwent MS/MS analyses. Mass and MS/MS spectra were analysed by the Masslynx software (Waters). The analysis was repeated on ten independently produced samples to test reproducibility.

2.3.2 Circular Dichroism analysis of HT-Pvfp-5β

Circular dichroism is a spectroscopic technique based on differential absorption of left and right handed circularly polarized light by molecules[84,85]. Not all molecules can be analyzed by this technique but only optically active ones or those who are chiral. The

concept of chirality is an important requirement for chiro-optical techniques such as CD. A molecule is chiral if it possesses a stereogenic centre, or if its spatial arrangement in space leads it to possess many chiral properties, such as proteins. The information available from a protein CD spectrum are related with protein secondary structure composition[86], such as percentage of α-helix, β-sheet, turns etc, or with protein stability, as the presence of folded and unfolded region.

In the present research work, CD studies were carried out with a J-815 (Jasco, Tokyo, Japan) spectropolarimeter using a quartz cell with 0.1 mm path length. Spectra were recorded at 20 °C, from 180 to 260 nm, using a scan rate of 100 nm/min a 1 nm bandwidth, and a data pitch of 0.5 nm. Each spectrum was the result of the average of 5 accumulations. The final concentration of the tested protein was 34 μM. All spectra were corrected by subtracting the spectra of the solvent. Three buffer conditions were tested: 5% acetic acid at pH 2.0, and sodium acetate buffer at 4.0 and 5.6 pH values.

2.3.4 Determination of Pvfp-5β structure in solution by NMR

Nuclear Magnetic Resonance (NMR) and X-ray crystallograph[87], along with the recent implemented cryo-EM technique[88], are the primary experimental methods used for three-dimensional protein structure determination. NMR offers the advantage of being able to determine the structure of proteins in solution, so giving information on the protein structure closer to a native situation than crystallography technique, where, instead, the protein is found in crystallized form[89–91]. However, NMR is limited by protein size and it can be applied typically to protein no bigger than 50–70 KDa. Pvfp-5β is a 9.5 KDa monomeric protein, thus NMR resulted very suitable for its structure characterization. In particular, structure of Pvfp-5β was determine at 298 K in 20 mM Sodium Acetate pH 4.5. Being NMR the unique technique intimately related to protein motions, dynamic processes of Pvfp-5β were also investigated.

For a physical point of view, NMR spectrometry exploits the absorption of electromagnetic radiation pulse, by protein's nuclei, which are dipped in a magnetic field. This process, known as perturbation, leads to nuclear spin transitions. When the perturbation ceases, nuclei return to the equilibrium in a process known as relaxation. The absorbed energy emitted during relaxation, is collected by the detector which produces

the *free induction decay* (FID), which is the sum, over time, of all the radiations emitted by the nuclei. The FID information, a function of time, is converted by Fourier transformation (FT) into a readable spectrum which is a function of frequency. NMR is routinely used for small molecules characterization, and nowadays it is possible to acquire and determine the structure of small molecules with almost full automated procedures. Protein structure determination by NMR requires, instead, several no automated steps, deep knowledge of the technique, and a tough and strict work is needed by users. Several bi-dimensional and three-dimensional NMR experiments have to be acquired, processed, and analysed in order to assign all resonances of the protein nuclei ^{1}H, ^{15}N and ^{13}C. These are then used to determine the distance between nuclei, and determine the overall structure of the protein by deep analysis of three-dimensional NOESY-HSQC spectra. These are based on the Nuclear Overhauser effect (NOE)[92] proposed in 1953 by Albert W. Overhauser. The NOE effect is the transfer of nuclear spin polarization from one nuclear spin population to another via dipolar cross-relaxation. NOESY-HSQC experiments allow to determine dipolar coupling, such as significant interprotonic distances, at $1/r^6$, that occur between atoms far in protein sequence but close in space, within 5-6 Ångström, due to protein folding (Fig. 2.4). The intensity of NOE peaks is directly related to protons distance, thus a list of distance restrains is generated, and it used as the main input in protein structure calculation softwares.

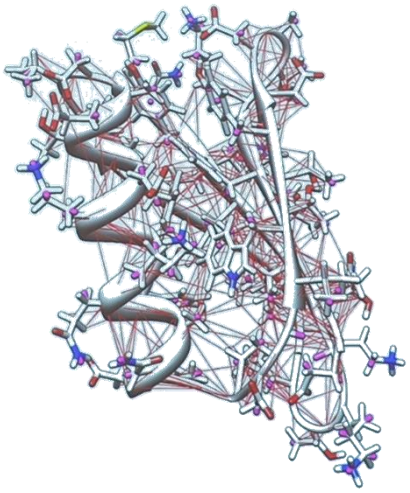


Figure 2.4 Nuclear Overhauser Effect. Structure of GB3 protein. Red lines show the distance restrains obtained by NOEs. Image taken from *CHIMIA*, 2012, **66**, 787-790

To determine Pvfp-5β structure, hetero-nuclear bidimensional ^{15}N-HSQC NMR experiments[93] were acquired on ^{15}N-labelled Pvfp-5β sample at concentration of approximately 280 μM using a Bruker AVANCE III TM HD NMR Spectrometer operating at 600 MHz (ATeN Center, University of Palermo). Three-dimensional triple-resonance NMR dataset for backbone assignment was carried out on ^{13}C,^{15}N-labelled Pvfp-5β using Bruker AVANCE III TM HD NMR Spectrometers operating at 800 MHz, available at the NMR Center of King's College London (London, UK), and at Ri.MED Foundation (Palermo, Italy). Two different samples were prepared for this triple resonance dataset, one at a concentration of approximately 350 μM and one at approximately 900 μM. Three-dimensional triple-resonance NMR dataset for side chains assignment and NOE restrains was acquired on ^{13}C,^{15}N-labelled Pvfp-5β using both a Bruker AVANCE III TM HD NMR Spectrometer operating at 700 MHz (MRC Biomedical NMR Centre - Crick Institute, London, UK), and a Bruker AVANCE III TM HD NMR Spectrometer operating at 800 MHz (Ri.MED Foundation, Palermo, Italy). Combined analysis of the triple resonance NMR experiments HNCACB[94], HN(CO)CaCb[95], HN(CO)CA[96] and HNCA[97] allowed full sequential assignment of Pvfp-5β backbone atoms. CO, Ha and Hb resonances were assigned by analysis of HNCO[98], HN(CO)HBHA and HNHBHA[99]. Side-chain resonances were assigned using aliphatic HCCH Total Correlation Spectroscopy (HCCH-TOCSY)[100], ^{15}N-NOESY-HSQC[101] and ^{13}C-NOESY-HSQC[102]. 2D (HB)CB(CGCD)HD, (HB)CB(CGCD)CEHE[103] and aromatic HCCH-COSY experiments[104] were acquired and analyzed for aromatic side chains resonance assignment. NOE data were collected from ^{15}N-NOESY HSQC and ^{13}C-NOESY HSQC[105]. The pulse sequence for each experiment is extensively detailed in ANNEX II.2, while experiments description and NMR parameters used in this research work are reported in the following sections. All spectra were processed using NMRPipe/NMRDraw software developed by F. Delaglio[106], and analyzed using CCPNMR software[107]. ARIA 2.3 software[108] was used for three-dimensional structure calculation.

^{1}H,^{15}N HSQC

To protein NMR experts, the ^{1}H,^{15}N HSQC (Heteronuclear Single Quantum Coherence) experiment[109–111] represents the fingerprint of the protein under observation, so it is the

first heteronuclear experiment generally acquired. It shows all the H-N correlations, such as the backbone amide protons (HN) of each amino acid present in the protein (except Proline), the Nε-Hε of side-chains of Tryptophan residues, and the Nδ-Hδ2/Nε-Hε2 of side-chains of Asparagine and Glutamine residues. In $^{1}H,^{15}N$ HSQC polarization is transferred from an amide proton to its own ^{15}NH nucleus via the J-coupling. The polarization is then transferred back to the proton, and signal recording follows (Fig. 2.5).

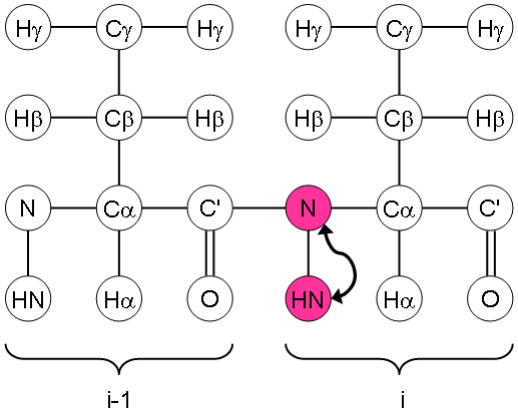


Figure 2.5. Magnetization transfer in 2D $^{1}H^{15}N$ HSQC experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

$^{1}H^{15}N$ HSQC spectrum for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, and in Echo-Antiecho mode for nitrogen indirect dimension. Carrier frequencies were centered at 4.7 ppm and 118 ppm for proton and nitrogen, respectively. 2048 points were collected for proton dimension (F2), and 128 points for nitrogen dimension (F1). The acquisition times were 0.0820 seconds for direct dimension and 0.0220 seconds for indirect dimension. The spectral width was 15.6 ppm for direct dimension, and 36 ppm for indirect dimension. 16 scans were acquired, resulting in 38 minutes for the total experiment.

$^{1}H,^{13}C$ HSQC

Bi-dimensional $^{1}H,^{13}C$ HSQC experiment[112] is the carbon equivalent of $^{1}H,^{15}N$ HSQC. It reveals the chemical shift of ^{1}H and ^{13}C nuclei directly bonded to each other via the ^{1}JCH coupling. The resulting spectrum appear in two dimensions, one for hydrogen and the second one for carbon. The magnetisation moves from ^{1}H to ^{13}C and then returns back for the detection (Fig. 2.6).

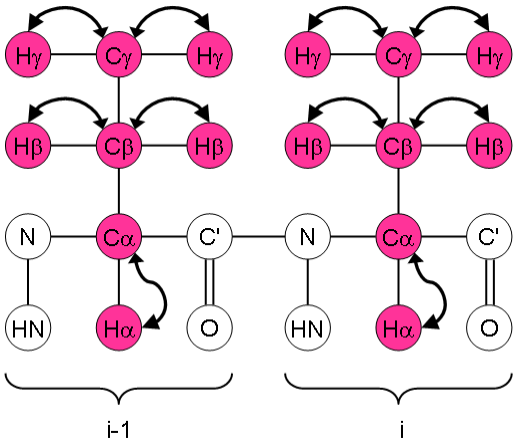


Figure 2.6. Magnetization transfer in 2D ^{1}H ^{13}C HSQC experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

^{1}H,^{13}C HSQC spectrum for ^{13}C,^{15}N-Pvfp-5β was acquired in DQD mode for proton direct dimension, and in Echo-Antiecho mode for carbon indirect dimension. Proton and carbon carrier frequencies were centered at 4.7 ppm, and 43 ppm, respectively. The number of points collected were 1024 for proton dimension (F2), and 512 for carbon dimension (F1). The acquisition time was 0.0409 seconds for proton direct dimension, and 0.0159 seconds for carbon indirect dimension. The spectral width was 15.6 ppm for proton dimension, and 80 ppm for carbon dimension. 16 scans were acquired, resulting in 1 hour and 47 minutes for the total experiment.

HNCACB

The three-dimensional HNCACB experiment[113], along with HN(CO)CACB, is useful for the assignment of the protein backbone atoms Cα and Cβ of each amino acid belonging to the protein sequence. In the corresponding spectrum, each amide group is correlated with both its own ^{13}Cα and ^{13}Cβ (intra-residue), and those of the preceding amino acid residue (inter-residue). In other words, for each NH group there are four visible peaks in the spectrum: two Cα and two Cβ. To obtain this, the magnetization is first transferred from ^{1}Hα and ^{1}Hβ to ^{13}Cα and ^{13}Cβ, respectively, and then from ^{13}Cβ to ^{13}Cα. It follows a transfer from ^{13}Cα to the amide nitrogen ^{15}NH, and then to the amide proton ^{1}HN for detection (Fig. 2.7). ^{13}Cα and ^{13}Cβ resonances of the preceding amino acid are observable because ^{15}N receives magnetization from both its own ^{13}Cα (residue *i*), and from the ^{13}Cα of the preceding residue (residue *i-1*).

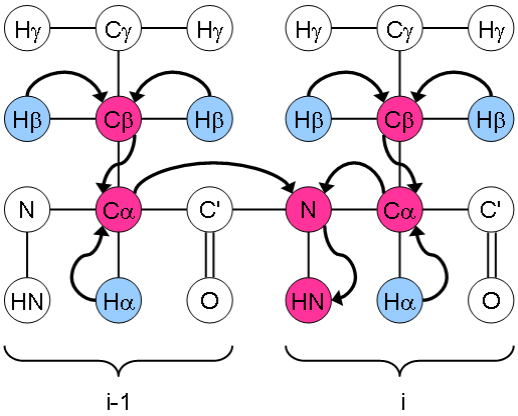


Figure 2.7. Magnetization transfer in 3D HNCACB experiment, Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HNCACB spectrum for ^{13}C,^{15}N-Pvfp-5β was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for nitrogen indirect dimension, and in States-TPPI mode for carbon indirect dimension. Carrier frequencies were centered at 4.7 ppm, 118 ppm, and 39 ppm for proton, nitrogen, and carbon, respectively. 1024 points were collected for proton dimension (F3), 60 points for nitrogen dimension (F2), and 120 points for carbon dimension (F1). The acquisition times were 0. 0.0450 seconds for proton direct dimension, 0.0102 seconds for nitrogen indirect dimension, 0.0040 seconds for carbon indirect dimension. The spectral width was 14.2 ppm for proton, 36 ppm for nitrogen, and 74 ppm for carbon. 8 scans were acquired, resulting in 18 hours and 37 minutes for the total experiment.

HN(CO)CACB

The three-dimensional HN(CO)CACB[114] is used in combination with the HNCACB experiment to assign the resonances of $^{13}C\alpha$ and $^{13}C\beta$. In the case of HN(CO)CACB, for each NH group there are two visible peaks in the spectrum corresponding to $^{13}C\alpha$ and $^{13}C\beta$ of the preceding residue (*i-1*). Comparing the two experiments is then possible to distinguish the resonances of the *i-1* residue from those of the *i* residue, and proceed with sequential assignment of the polypeptide chain. In HN(CO)CACB, the first polarization transfers involve atoms of the *i-1* residue. In particular, polarization is transferred from $^{1}H\alpha$ and $^{1}H\beta$ to $^{13}C\alpha$ and $^{13}C\beta$, respectively, then from $^{13}C\beta$ to $^{13}C\alpha$, and as last from to $^{13}C\alpha$ to ^{13}CO. At this point, the magnetization is transferred to ^{15}NH of residue *i*, and then to its amide proton ^{1}HN, and signal recording follows (Fig. 2.8). The resulting spectrum

occur in one dimension for both $^{13}C\alpha$ and $^{13}C\beta$, and for ^{15}NH and ^{1}HN in the others two dimensions.

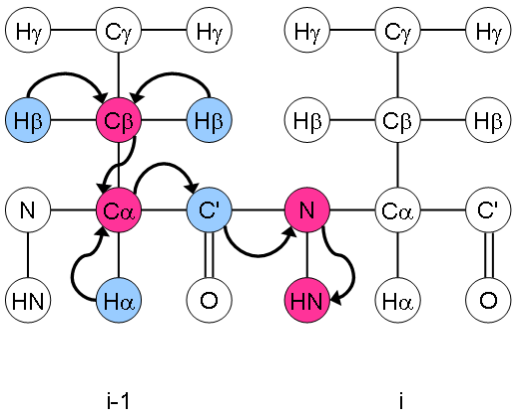


Figure 2.8. Magnetization transfer in 3D HN(CO)CACB experiment, Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HN(CO)CACB spectrum for ^{13}C,^{15}N-Pvfp-5β was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for nitrogen indirect dimension, and in States-TPPI mode for carbon indirect dimension. Carrier frequencies were centered at 4.7 ppm, 118 ppm, and 39 ppm for proton, nitrogen, and carbon, respectively. 1024 points were collected for proton dimension (F3), 60 points for nitrogen dimension (F2), and 120 points for carbon dimension (F1). The acquisition times were 0. 0.0450 seconds for proton direct dimension, 0.0102 seconds for nitrogen indirect dimension, 0.0040 seconds for carbon indirect dimension. The spectral width was 14.2 ppm for proton, 36 ppm for nitrogen, and 74 ppm for carbon. 8 scans were acquired, resulting in 18 hours and 51 minutes for the total experiment.

HN(CO)CA

HN(CO)CA experiment[115] is used in conjunction with the HNCA experiment in order to distinguish between its own $^{13}C\alpha$ and those of the previous residue *i-1*. This is only selective for $^{13}C\alpha$ of the preceding residue *i-1*. The magnetisation moves from the amide proton ^{1}HN of residue *i* to its related amide nitrogen ^{15}NH, and then is transferred to ^{13}CO of the preceding residue *i-1*. From here it is transferred to $^{13}C\alpha$ of residue *i-1* and the chemical shift is evolved. The polarization then returns the same way to ^{1}HN of residue *i* for detection (Fig. 2.9). The chemical shift evolves for the amide proton, the nitrogen and carbon alpha, but it does not evolve for carbonyl carbon.

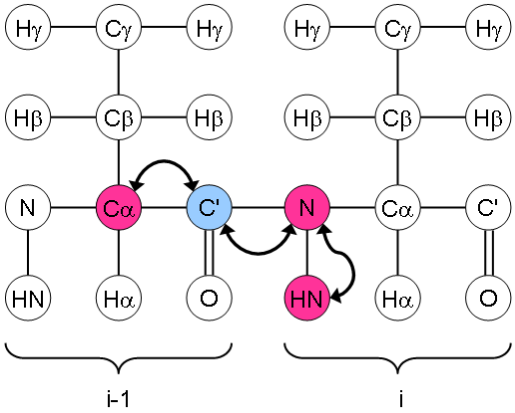


Figure 2.9. Magnetization transfer in 3D HN(CO)CA experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HN(CO)CA experiment for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, in States-TPPI mode for nitrogen indirect dimension, and in States-TPPI mode for carbon indirect dimension. Proton, nitrogen and carbon carrier frequencies were centered at 4.7 ppm, 116 ppm, and 53.2 ppm, respectively. The number of points collected were 2048 for proton dimension (F3), 40 for nitrogen dimension (F2), and 128 for carbon dimension (F1). The acquisition time was 0.0819 sec for proton, 0.0082 sec for nitrogen, and 0.0106 for carbon. The spectral width was 15.6 ppm for proton, 30 ppm for nitrogen, and 30 ppm for carbon. 8 scans were acquired, resulting in 13 hours and 39 minutes for the total experiment.

HNCA

Like the HN(CO)CA, HNCA experiment[98] is used to determine ^{13}Cα resonances. Each amide nitrogen is coupled both to the ^{13}Cα of its own residue *i*, and to the ^{13}Cα of the preceding residue *i-1*. As consequence, in the corresponding spectrum two peaks are visible, one for the ^{13}Cα of the preceding residue *i-1*, and one more intense for the ^{13}Cα of residue *i*. The polarization moves from ^{1}HN to ^{15}NH, and to the ^{13}Cα. It then returns back to nitrogen and finally to ^{1}HN for signal recording (Fig. 2.10). The chemical shift is evolved for ^{1}HN as well as for ^{15}NH and ^{13}Cα.

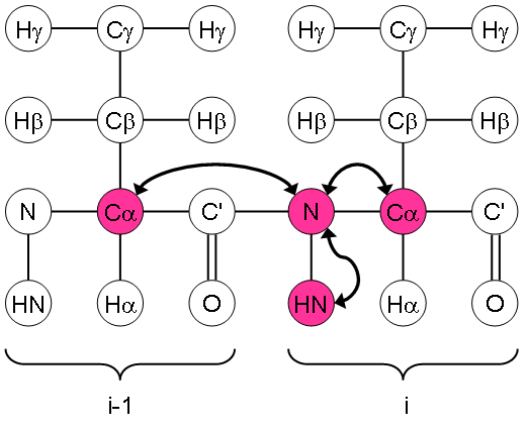


Figure 2.10. Magnetization transfer in 3D HNCA experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HNCA experiment for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, and in States-TPPI mode for both nitrogen and carbon indirect dimensions. Proton, nitrogen and carbon carrier frequencies were centered at 4.7 ppm, 116 ppm, and 53.2 ppm, respectively. The number of points collected were 2048 for proton dimension (F3), 48 for nitrogen dimension (F2) and 128 for carbon dimension (F1). The acquisition time was 0.0901 seconds for proton, 0.0098 seconds for nitrogen, and 0.0106 seconds for carbon. The spectral width was 14.2 ppm for proton, 30 ppm for nitrogen, and 30 ppm for carbon. 8 scans were acquired, resulting in 16 hours and 15 minutes for the total experiment.

HNCO

HNCO experiment[116] is used to assign the carbonyl carbon resonances. Polarization is first transferred from ^{1}HN to ^{15}NH of residue *i*, then it goes to the directly attached carbonyl ^{13}CO of residue *i-1* by the $^{15}NH–^{13}CO$ J-coupling. It then returns the same way to ^{1}HN for detection (Fig. 2.11). All the three involved nuclei are detected.

HNCO spectrum for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for nitrogen indirect dimension and in States-TPPI mode for carbon indirect dimension. Proton, nitrogen and carbon carrier frequencies were centered at 4.7 ppm, 118 ppm, and 173.5 ppm, respectively. The number of points collected were 1024 for proton dimension (F3), 40 for nitrogen dimension (F2) and 120 for carbon dimension (F1). The acquisition time were 0.0450 seconds for proton, 0.0068

seconds for nitrogen, and 0.0213 seconds for carbon. The spectral width was 14.2 ppm for proton, 36 ppm for nitrogen, and 14 ppm for carbon. 8 scans were acquired, resulting in 12 hours and 21 minutes for the total experiment.

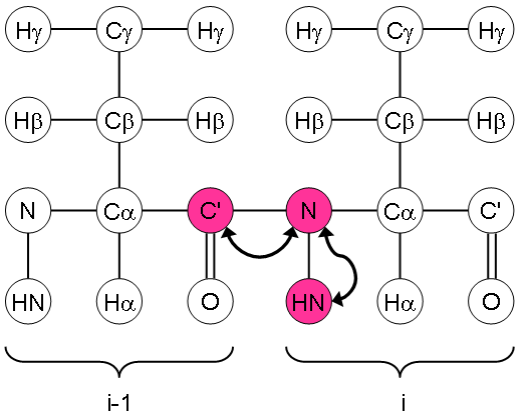


Figure 2.11. Magnetization transfer in 3D HNCO experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HBHA(CO)NH

The three-dimensional HBHA(CO)NH[117] is designed to correlate the amide proton ^{1}HN of a residue *i* to ^{1}Hα and ^{1}Hβ of its preceding residue *i-1*. The magnetisation transfer firstly involves only nuclei of the residue *i-1*. In particular, it is transferred from ^{1}Hα and ^{1}Hβ to ^{13}Cα and ^{13}Cβ, respectively, and then from ^{13}Cβ to ^{13}Cα, and from ^{13}Cα to ^{13}CO. At this point, the magnetization is transferred to the amide nitrogen ^{15}NH of residue *i*, and afterwards to its own ^{1}HN for detection (Fig. 2.12). In the resulting spectrum, the chemical shift is evolved in one dimension for both ^{1}Hα and ^{1}Hβ, and for ^{15}NH and ^{1}HN in the others two dimensions.

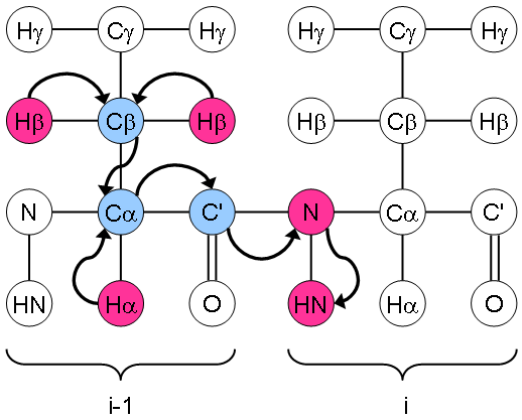


Figure 2.12. Magnetization transfer in 3D HBHA(CO)NH experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HBHA(CO)NH spectrum for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for nitrogen indirect dimension, and in States-TPPI mode for proton indirect dimension. Proton and nitrogen carrier frequencies were centered at 4.7 ppm and 118 ppm, respectively. The number of points collected were 1024 for direct proton dimension (F3), 40 for indirect nitrogen dimension (F2) and 80 for indirect proton dimension (F1). The acquisition time were 0.0409 seconds for proton direct dimension, 0.0068 seconds for nitrogen, and 0.0062 seconds for proton indirect dimension. Spectral widths were set at 15.6 ppm, 35.9 ppm, and 8.0 ppm, for direct proton, nitrogen, and indirect proton, respectively. 16 scans were acquired, resulting in a total experiment time of 16 hours and 53 minutes.

HBHANH

HBHANH experiment[118] is used in conjunction with HBHA(CO)NH to assign the resonances of $^{1}H\alpha$ and $^{1}H\beta$. Unlike the HBHA(CO)NH experiment, the magnetisation is not transferred through the ^{13}CO, and each amide group is correlated with both its own $^{1}H\alpha$ and $^{1}H\beta$, and those of the preceding residue *i-1* (Fig. 2.13). As seen for HBHA(CO)NH, in HBHANH spectrum the chemical shift is evolved in one dimension for both $^{1}H\alpha$ and $^{1}H\beta$, and for ^{15}NH and ^{1}HN in the others two dimensions.

HBHANH spectrum for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for nitrogen indirect dimension, and in States-TPPI mode for proton indirect dimension. Proton and nitrogen carrier frequencies were centered at 4.7 ppm, and 118 ppm, respectively. The number of points collected were 1024 for direct proton dimension (F3), 40 for indirect nitrogen dimension (F2), and 80 for indirect proton dimension (F1). The acquisition time were 0.0409 seconds for proton direct dimension, 0.0068 seconds for nitrogen, and 0.0062 seconds for proton indirect dimension. The spectral width was 15.6 ppm for direct proton, 35.9 ppm for nitrogen, and 8 ppm for indirect proton. 16 scans were acquired, resulting in 16 hours and 43 minutes for the total experiment.

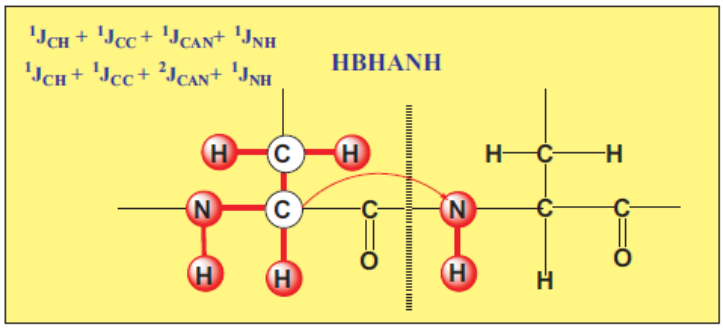


Figure 2.13. Magnetization transfer in 3D HBHANH experiment. Image taken from http://biotool.xmu.edu.cn/guide/eNMR/eNMR3Dprot/hbhanh.html

HCCH-TOCSY

The three-dimensional HCCH-TOCSY experiment[119,120] is mostly useful for side chain resonances assignment, and it is specially designed to correlate all side-chain aliphatic protons and ^{13}C resonances through $^1J(CH)$ and $^1J(CC)$ coupling constants. Polarization moves from the side-chain or backbone hydrogen nuclei to their attached ^{13}C nuclei. It follows the isotropic ^{13}C mixing, and then polarization returns back to the side-chain hydrogen atoms for detection (Fig. 2.14). The length of mixing time determines how far the magnetization is transferred. Transfer either occurs for aliphatic ^{13}C nuclei or for aromatic ^{13}C nuclei depending on the ^{13}C carrier frequency used during the acquisition.

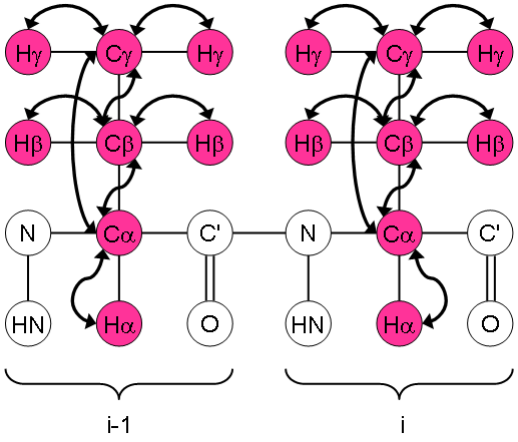


Figure 2.14. Magnetization transfer in 3D HCCH-TOCSY experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

HCCH-TOCSY spectrum for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, and in States-TPPI mode for both carbon indirect dimension, and proton indirect dimension. Proton and carbon carrier frequencies were centered at 4.7 ppm, and 42 ppm, respectively. The number of points collected were 2048 for direct proton dimension (F3), 128 for indirect carbon dimension (F2) and 256 for indirect proton dimension (F1). The acquisition time was 0.0819 seconds for proton direct dimension, 0.0049 seconds for carbon, and 0.026 seconds for proton indirect dimension. The spectral width was 15.6 ppm for direct proton, 63.9 ppm for carbon, and 6 ppm for indirect proton. 4 scans were acquired, resulting in 42 hours and 49 minutes for the total experiment.

HCCH-COSY

A complementary experiment to HCCH-TOCSY is the HCCH-COSY (Correlated Spectroscopy) experiment[121], which is based on transfer over one ^{13}C-^{13}C bond and correlates neighbouring ^{1}H frequencies only.

HCCH-COSY spectrum for $^{13}C,^{15}N$-Pvfp-5β was acquired in DQD mode for proton direct dimension, and in States-TPPI mode for both carbon and proton indirect dimensions. Proton and carbon carrier frequencies were centered at 4.7 ppm, and 126 ppm, respectively. The number of points collected were 4096 for direct proton dimension (F3), 50 for indirect carbon dimension (F2) and 200 for indirect proton dimension (F1). The acquisition time was 0.180 seconds for proton direct dimension, 0.0044 seconds for carbon, and 0.010 seconds for proton indirect dimension. The spectral width was 14.2 ppm for direct proton, 28 ppm for carbon, and 12 ppm for indirect proton. 8 scans were acquired, resulting in 27 hours and 59 minutes for the total experiment.

(HB)CB(CGCD)HD and (HB)CB(CGCD)CEHE

Bi-dimensional (HB)CB(CGCD)HD and (HB)CB(CGCD)CEHE experiments[122] are particularly designed to provide correlations between carbon $^{13}C\beta$ and protons $^{1}H\delta$ and $^{1}H\varepsilon$ of aromatic rings. The approach is based exclusively on the transfer of magnetization via scalar couplings. The path of magnetization transfer in (HB)CB(CGCD)HD experiment starts from $^{1}H\beta$ to $^{13}C\beta$, and then moves to $^{13}C\gamma$ and $^{13}C\delta$ to detect $^{1}H\delta$. In

(HB)CB(CGCD)CEHE experiment, the magnetization moves from ^{1}Hβ to ^{13}Cβ, from ^{13}Cβ to ^{13}Cγ and ^{13}Cδ, and finally it moves to ^{13}Cε to detect ^{1}Hε (Fig. 2.15).

(HB)CB(CGCD)HD and (HB)CB(CGCD)CEHE spectra for ^{13}C,^{15}N-Pvfp-5β were acquired in DQD mode for proton direct dimension, and in States-TPPI mode for carbon indirect dimension. Proton and carbon carrier frequencies were centered at 4.7 ppm, and 32.0 ppm, respectively. The number of points collected were 4096 for proton dimension (F2), and 256 for carbon dimension (F1). The acquisition time was 0.2129 seconds for direct dimension, and 0.0318 seconds for indirect dimension. The spectral width was 12 ppm for direct dimension, 20 ppm for the indirect one. 80 scans were acquired, resulting in 1 hour and 54 minutes of acquisition time for each experiment.

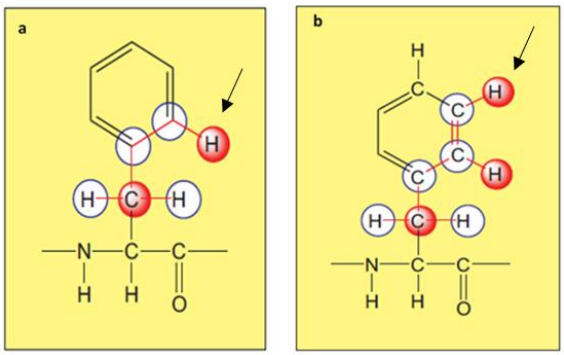


Figure 2.15. Magnetization transfer in bi-dimensional (HB)CB(CGCD)HD[a] and (HB)CB(CGCD)CEHE[b] experiments. Image taken from http://triton.iqfr.csic.es/guide/eNMR/eNMRtri/hbcbhd.html

^{15}N NOESY-HSQC

The three-dimensional ^{15}N NOESY (Nuclear Overhauser Effect Spectroscopy)-HSQC experiment[123–125] is, along with the ^{13}C NOESY HSQC, the most useful method for protein structure determination. The experiment consists of a ^{1}H-^{15}N HSQC part, where signals are dispersed according to the backbone amide protons frequency, and a ^{1}H-^{1}H NOESY part, that contains NOE restraints between NH and all the spatially neighbouring protons either within the same or in other residues. The spectrum is characterized by the presence of hydrogen in two dimensions, and nitrogen in the third dimension, and it appears with a diagonal where amide protons of each residue fall, and cross peaks related to spatial information. The magnetisation is exchanged between all hydrogens using the NOE, and then is transferred to neighbouring ^{15}N nuclei and back to ^{1}H for detection (Fig. 2.16). The amplitude of magnetization propagation is related to the mixing time value.

^{15}N NOESY-HSQC spectrum for ^{13}C,^{15}N-Pvfp-5β was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for nitrogen indirect dimension, and in States-TPPI mode for proton indirect dimension. Proton and nitrogen carrier frequencies were centered at 4.7 ppm, and 116 ppm, respectively. The number of points collected were 2048 for direct proton dimension (F3), 60 for indirect nitrogen dimension (F2) and 160 for indirect proton dimension (F1). The acquisition time was 0.0106 seconds for proton direct dimension, 0.0012 seconds for nitrogen, and 0.0083 seconds for proton indirect dimension. The spectral width was 12 ppm for direct dimension, 32 ppm for nitrogen, and 12.0 ppm for proton indirect dimension. 8 scans were acquired, resulting in 27 hour and 22 minutes of acquisition time for the entire experiment.

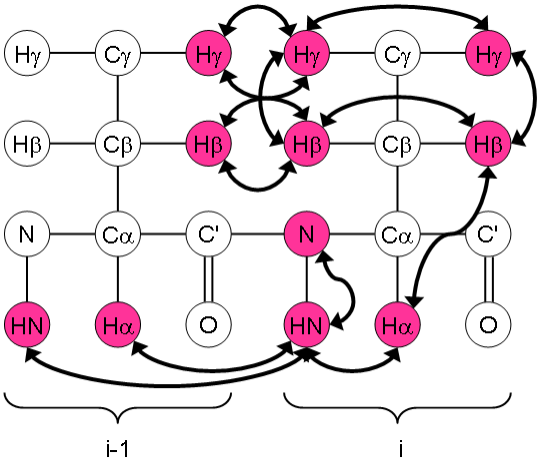


Figure 2.16. Magnetization transfer in 3D ^{15}N NOESY-HSQC experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

^{13}C NOESY-HSQC

Analogously to the ^{15}N NOESY HSQC experiment, the three-dimensional ^{13}C NOESY HSQC [123–125] is also used to correlate hydrogens that are close in space and thus to achieve protein structure information. Magnetization transfer either occurs to and from the aliphatic ^{13}C nuclei or to and from the aromatic ^{13}C nuclei depending on the ^{13}C carrier frequency used during the acquisition. Polarization moves through all hydrogens using the NOE, then it is transferred to close ^{13}C nuclei, and finally it returns back to ^{1}H for detection (Fig. 2.17). The amplitude of magnetization propagation is related to the mixing time value.

^{13}C NOESY-HSQC spectrum for ^{13}C,^{15}N-Pvfp-5β related to aliphatic side-chains nuclei was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for carbon indirect dimension, and in States-TPPI mode for proton indirect dimension. Proton and carbon carrier frequencies were centered at 4.7 ppm, and 42 ppm, respectively. The number of points collected were 1024 for direct proton dimension (F3), 80 for indirect carbon dimension (F2) and 220 for indirect proton dimension (F1). The acquisition time was 0.053 seconds for proton direct dimension, 0.0031 seconds for carbon indirect dimension, and 0.0015 seconds for proton indirect dimension. The spectral width was 12 ppm for direct dimension, 64 ppm for carbon, and 9 ppm for proton indirect dimension. 16 scans were acquired, resulting in 3 days, 23 hours and 41 minutes of acquisition time for the entire experiment.

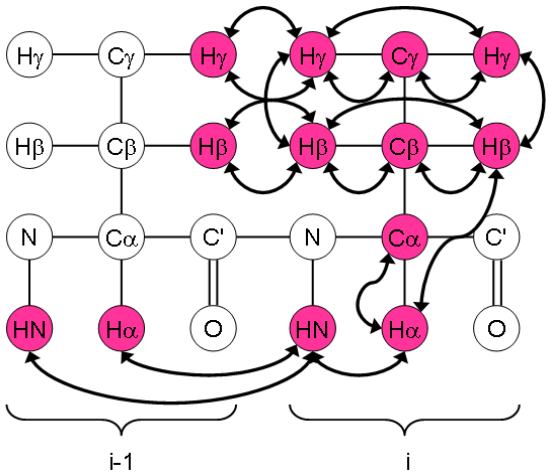


Figure 2.17. Magnetization transfer in 3D ^{13}C NOESY-HSQC experiment. Image from Protein NMR | A Practical Guide (protein-nmr.org.uk).

^{13}C NOESY-HSQC spectrum for ^{13}C,^{15}N-Pvfp-5β related to aromatic side-chains nuclei was acquired in DQD mode for proton direct dimension, in Echo-Antiecho mode for carbon indirect dimension, and in States-TPPI mode for proton indirect dimension. Proton and carbon carrier frequencies were centered at 4.7 ppm, and 126 ppm, respectively. The number of points collected were 4096 for direct proton dimension (F3), 80 for indirect carbon dimension (F2), and 180 for indirect proton dimension (F1). The acquisition time was 0.2130 seconds for proton direct dimension, 0.007 seconds for carbon, and 0.0012 seconds for proton indirect dimension. The spectral width was 12 ppm for direct dimension, 28 ppm for carbon, and 9 ppm for proton indirect dimension. 8 scans were acquired, resulting in 44 hours and 6 minutes of acquisition time for the entire experiment.

2.4 Molecular Dynamic studies on Pvfp-5β through NMR spectroscopy

Think of molecules as rigid objects that simply rotate or translate within the solution through linear diffusion is an incorrect conclusion. The more complex a molecule is, the more regions could exist that move independently to each other, to which a specific function is linked. This results very important for protein function that could be attribute to the combination of structure and molecular motions. Backbone or side-chain motions are characterized by low energy barriers, and thus they occur on very fast time scales, such as at pico- and nano-seconds. Motions with higher energy barriers occur on longer time scales, such as µs-ms for catalysis, ligand binding, and allostery, and ms-s for unfolding events and complex formation between multiple components[126] (Fig. 2.18).

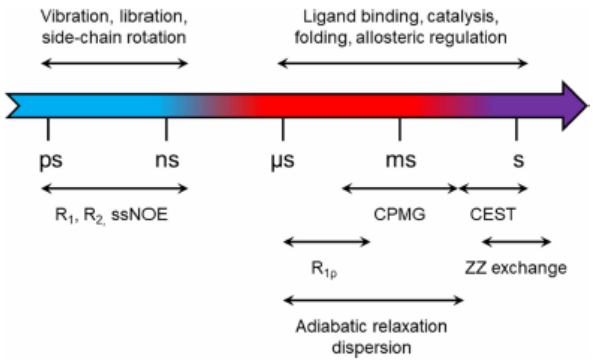


Figure 2.18. Correlation of molecular motions with NMR observables. Summary of the timescales of biomolecular dynamics and the major NMR relaxation techniques. Blue: ps–ns; red: µs–ms; purple: ms–s. Image taken from *Emerging Topics in Life Sciences,* 2018, 2, 93-105

NMR is a technique particularly well suited to address protein motions and fluctuations through nuclei relaxation[127–135]. This is the process by which spins in a molecule come to equilibrium with the surroundings after a perturbation. In more details, if the equilibrium state is perturbated by applying a radio frequency of 90 degrees pulse, the population of the energy levels is altered, generating a transverse magnetization, and phase coherence is induced to the nuclei. When the perturbation ceases, the magnetisation returns to the equilibrium by transitions between the two energy levels. This process results in relaxation. The relaxation spin-reticulum or longitudinal relaxation time, T1, is related to the time required to re-establish thermal equilibrium through the come back of nuclei

from high energy state to a low energy state releasing energy back to the surrounding lattice. The spin-spin relaxation or transverse relaxation time T2, is related to dephasing of the transverse component of the magnetisation that characterizes the exponential decay of the signal. The two relaxation mechanisms by which nuclei with spin ½, the most interesting in biomolecular field, relax are dipole-dipole coupling and chemical shift anisotropy. The frenetic and fast movement with which molecules move in solution, causes the induced fields, produced from both the two interactions, in turn to move and create local fluctuations of the magnetic field leading the total field to slightly fluctuates too and to be felt by nuclear spin in the molecules. This wobbling of the local magnetic field is random due to the random molecular motion, thus the field felt by the spin will also vary randomly over time. Describing a random motion of a molecule or a molecular region is possible thanks to a temporal evolution of a function G(τ) (Eq. 2.1) which describe the magnetic field felt at time τ by a spin present on the molecule or a molecular region.

(Eq.2.1) $G(\tau) = \langle Bx(t)\, Bx(t+\tau) \rangle$

The Eq.2.1 is known as *autocorrelation function* and describe how rapidly the magnetic field fluctuates. This function will be large and positive for small values of τ, as the motion will have move little the molecules from the initial position thus little altered the field, but will tend to assume values random, positive or negative, for large values of τ, when memory of the molecules initial position will be lost. Thus, the autocorrelation function could be seen as a memory function of the initial position of molecules and tends to be large for small values of τ and tends to zero for large value of τ (Fig. 2.19).

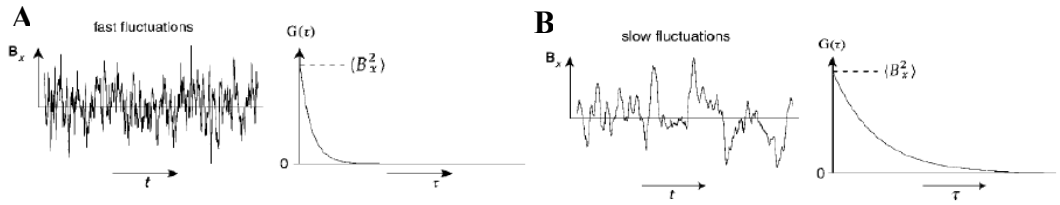


Figure 2.19. (A) Fast field fluctuations led to a fast decay of autocorrelation function for small value of τ. (B) Slow field fluctuations led to a slow decay of autocorrelation function for high value of τ.

Since the decay function is exponential the autocorrelation function takes the form of an exponential (Eq.2.2).

(Eq.2.2) $G(\tau) = \langle B^2x \rangle\, e^{-|\tau|/\tau c}$

In (Eq.2.2) τc is defined *as correlation time* of the fluctuations and thus is a measure of how fast the motion occurs. When the Fourier Transformation is applied to the correlation function (Eq.2.2), the *Spectral density* function is obtained, which describe any relaxation processes, involving a rotation at various frequencies ω (Eq.2.3).

(Eq.2.3) $$J(\omega) = 2 \langle B^2 x \rangle \tau c / 1+\omega^2 \tau^2 c$$

The probability that a transition could be achieved, relating to a fluctuant field, in dipole-dipole coupling depends on the type of spins gyromagnetic ratio (γ), on their spatial distance, the closer they are the stronger will be the fluctuating field, on their relative motion. To study protein molecular dynamic is used 15N as the dynamic probe because it presents an excellent dispersion, even in partially folded proteins, thus is possible to measure the local relaxation of each protein residues. The mechanisms by which the ^{15}N relaxation occurs are dipole-dipole coupling and chemical shift anisotropy. The ^{15}N dipole-dipole coupling is completely dominated by its attached proton(s) and this coupling is related to distance. Since there is only one proton closer to ^{15}N nuclei, around 1 Å, it could be assessed that this proton completely rules ^{15}N relaxation. The distance between proton and nitrogen is fixed and it does not depend on molecules conformation and does not change over time. The chemical shift anisotropy relaxation is characteristic only by ^{15}N nuclei and thus is only related to its $J(\omega_N)$ and also depends on an anisotropy factor. Laws regulating amide nitrogen nuclei relaxation are reported below (Eq.2.4-2.6).

(Eq.2.4)

$$1/T1= d^2[J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H+\omega_N)] + c^2 J(\omega_N)$$

(Eq.2.5)

$$1/T2= 0.5d^2[4J(0) + J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H)+6J(\omega_H-\omega_N)] +1/6\ c^2\ [3J(\omega_N) + 4J(0)]$$

(Eq.2.6)

$$NOE= 1 + T1\ (\gamma_H/\ \gamma_N)d^2\ [6J(\omega_H+\omega_N)+4J(\omega_H-\omega_N)]$$

Relaxation times, T1 and T2, depends on constant terms and on frequency terms that are unknown and related to protein motions. NOE equation is related to proton and nitrogen and depend on T1 while d^2 and c^2 are constant terms and are related to proton-nitrogen distance and to anisotropy factor respectively. $J(\omega_H)$, $J(\omega_H + \omega_N)$ and $J(\omega_H - \omega_N)$ behave

very similar thus they can be replaced by one J term by approximation. Finally, T1,T2 and heteronuclear NOE, can be obtained experimentally.

T_1, T_2, and heteronuclear NOE of Pvfp-5β were measured through bi-dimensional experiments on Bruker AVANCE III TM HD, operating at 800 MHz and 600 MHz at 298°K, located at Ri.MED Foundation and Aten Center, respectively.

T1 relaxation rates measurement

Pseudo three-dimensional $^{1}H,^{15}N$ HSQC spectra, for longitudinal relaxation time T1 of Pvfp-5β were acquired in QF (no frequency) mode for proton dimension, and in Echo-Antiecho mode for nitrogen dimension. Proton and nitrogen carrier frequencies were centered at 4.7 ppm, and 116 ppm, respectively. The number of points collected were 12 for proton dimension (F1), and 128 for nitrogen dimension (F2). The acquisition time was 0.00125 seconds for proton, and 0.0247 seconds for nitrogen. The number of scans acquired were 16. The spectral width was 12 ppm for direct dimension, and 32 ppm for indirect dimension. ^{15}N T1 values were measured from the spectra recorded with ten different durations of the delay T: T = 0.01, 0.06, 0.1, 0.2, 0.4, 0.6, 0.8, 1.2, 1.5 and 2.0 seconds. The experiment was recorded for 2 days and 1 hour.

T2 relaxation rates measurement

Pseudo three-dimensional $^{1}H,^{15}N$ HSQC spectra, for transverse relaxation time T2 of Pvfp-5β were acquired in QF (no frequency) mode for proton dimension, and in Echo-Antiecho mode for nitrogen dimension. Proton and nitrogen carrier frequencies were centered at 4.7 ppm, and 116 ppm, respectively. The number of points collected were 10 for proton dimension (F1) and 128 for nitrogen dimension (F2). The acquisition time was 0.00104 seconds for proton dimension and 0.0247 seconds for nitrogen dimension. The number of scans acquired were 48. The spectral width was 12 ppm for direct dimension and 32 ppm for indirect dimension. ^{15}N T2 values were measured from the spectra recorded with eight different durations of the delay T: T = 0.017, 0.034, 0.068, 0.135, 0.170, 0.203, 0.237, and 0.271 seconds. The experiment was recorded for 1 day and 1 hour.

T1 and T2 values were extracted by measuring peak intensities in 2D spectra as a function of a relaxation delay.

H-N NOE measurement

Heteronuclear $^{15}N,^{1}H$ NOEs spectra for Pvfp-5β were acquired in QF (no frequency) mode for proton dimension, and in Echo-Antiecho mode for nitrogen dimension. Proton and nitrogen carrier frequencies were centered at 4.7 ppm, and 116 ppm, respectively. The number of points collected were 2 for proton dimension (F1), and 128 for nitrogen dimension (F2). The acquisition time was 0.0104 seconds for proton dimension, and 0.0247 seconds for nitrogen dimension. The number of scans acquired were 80. The spectral width was 12 ppm for direct dimension and 32 ppm for indirect dimension. The experiment was recorded for 1 day and 22 hours. Heteronuclear NOEs were measured from the ratio of cross-peak intensities in the two experiments: with and without saturation of HN resonances.

The T1, T2 and heteronuclear NOE data sets were processed with Dynamic center software [136].

3. Results and Discussion

3.1 Bioinformatic studies of Pvfp-5β

Protein sequence carry within them very useful informations about protein structure and function. Each amino acid assumes a specific and not random position to which a characteristic function is linked, be it biological or structural. As consequence, probing protein sequences similarity is very useful to identify similarity among function, structure, or evolutionary relationships between proteins.

In order to identify structure and function similarities between Pvfp-5β and proteins already known in Protein Data Bank (www.rcsb.org), its amino acid sequence was analyzed through the bioinformatic tool BLAST[141]. This identified two tandem EGF motifs along the Pvfp-5β sequence that shares 47–50% identity with the EGF repeats of the Notch ligand delta-like 1 protein[142] (residues 286–368, PDB 4XBM, Fig. 3.1 A). The Notch protein is a transmembrane receptor with repeated extracellular EGF domains, and it is involved in many biological functions, such as angiogenesis, wound healing, bond and tissue growth. On a structural point of view, each EGF domain is composed by two β-sheet stabilized by three disulfide bridges, which may link calcium ions (Fig. 3.1 B). The high sequence similarity found among Pvfp-5β and tandem EGF domains in Notch ligand delta-like 1 protein, suggests a tight similarity in the folding, also confirmed by I_TASSER[143] secondary structure prediction (Fig. 3.2). Moreover, the presence of EGF domains in *mfps* is not unique of Pvfp-5β since other mussel adhesive proteins of various species, such as *Mgfp-2* of *Mytilus galloprovincialis*, have several copies of EGF-like motifs.

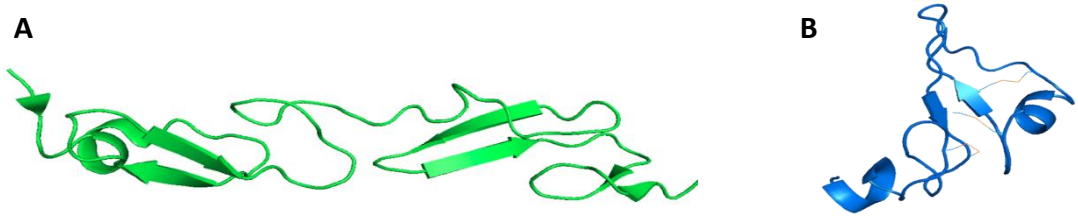


Figure 3.1. **A.** Notch ligand delta-like 1 protein, residues 286–368 (PDB 4XBM). **B.** EGF domain (PDB 2KV4)

Secondary structure prediction of Pvfp-5β

```
                    20                  40                  60                  80
                    |                   |                   |                   |
Sequence    VYYPNPCSPYPCRNGGTCKKRGLYSYKCYCRKGYTGKNCQYNACFPNPCLNGGTCGYVYGYPYYKCSCPYGYYGKQCQLKKY
Prediction  CCCCCCCCCCCCCCCCSSSSCCCCCSSSSCCCCCCCCCCCCCCCCCCCCCCCCCSSSCCCCCCCSSSSCCCCCCCCCCCCCC
Conf.Score  9575558798889996985089999887697998178874686669888999891066899998858938980786705879
            H:Helix; S:Strand; C:Coil
```

Figure 3.2. I_TASSER secondary structure prediction for Pvfp-5β

A structural model was then built submitting the alignment of Pvfp-5b with Notch ligand delta-like 1 protein to the SWISSMODEL server[144]. In the resulting model, the disulfide bridge between Cys8 and Cys19, in the N-terminal of HT-Pvfp-5β, is not formed probably because it is to close to the fraying edge of the domain. Moreover, the conformation of the linker between the two EGF motifs, which contains Cys45, does not allow the formation of the disulfide bridge Cys45–Cys56 as the consequence of a two-residue deletion with respect Notch sequence (Fig. 3.3). To better identify the disulfide bridges in Pvfp-5β, an experimental approach by mass spectrometry was later performed (paragraph 3.4).

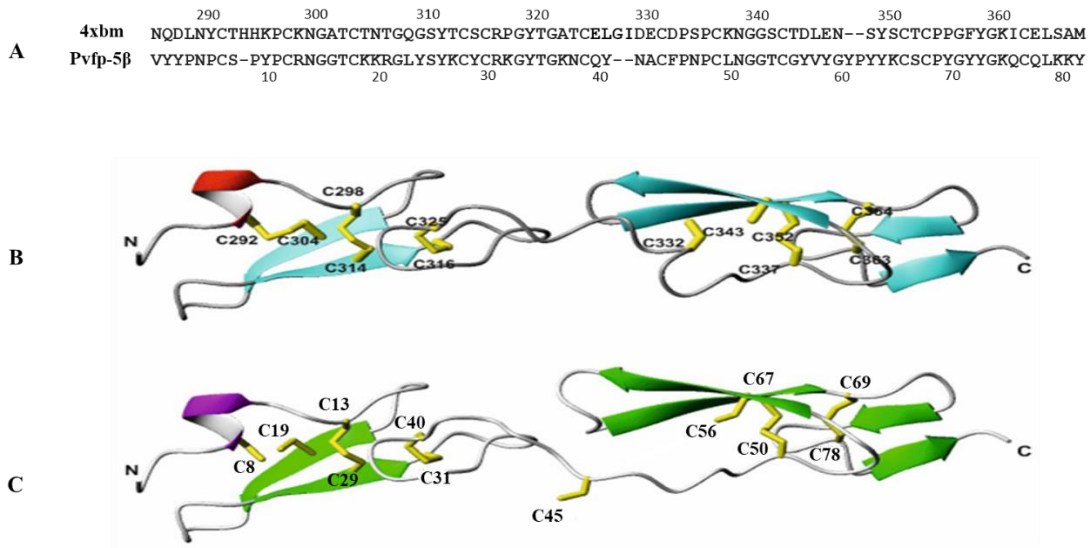


Figure 3.3. Structural model of HT-Pvfp-5β based on sequence similarity. **A.** Sequence alignment between Pvfp-5β (top) and the Notch ligand delta-like 1 protein (bottom). All the twelve cysteines are conserved. **B.** The crystal structure of Notch ligand delta-like 1 protein (PDB 4XBM). The disulfide bridges observed in the Notch structure are indicated. **C.** Comparative model of Pvfp-5β using 4XBM as a structural template.

3.2 Production of recombinant Pvfp-5β

Several attempts were done to optimize the production of recombinant Pvfp-5β as this was the prerequisite of any structural characterization of this protein. Pvfp-5β contains 12 cysteine residues in its primary sequence, it was then expected that it could be expressed into cell inclusion bodies. To increase protein solubility and avoid exhausting steps of extraction and refolding, Pvfp-5β was first expressed as a thioredoxin fusion protein with a tobacco etch virus (TEV) cleavable site just before the Pvfp-5β (protein sequence and parameters in ANNEX III.1). Despite the presence of thioredoxin, the fused protein was expressed in inclusion bodies. The construct was then modified and subcloned as a TEV protease cleavable N-terminal His-tag protein (HT-Pvfp-5β), which, although still resulting in inclusion bodies, allowed to attempt refolding of the only Pvfp-5β. The cloning strategy used was that of amplifying by PCR the entire plasmid except the region encoding for thioredoxin. This was obtained using a pair of mutually complementary primers annealing to the DNA regions flanking the thioredoxin coding. Deletion was confirmed through Sanger DNA sequencing (sequence results in ANNEX III.2). Expression tests of HT-Pvfp-5β were then performed in three different *E.coli* cell strains, two of which engineered to increasing the yield of the expressed protein in the soluble fraction (ArcticExpress and Origami). Unfortunately, all tested conditions led protein to inclusion bodies (Fig. 3.4), so it was necessary to implement a procedure to both extract the protein from the insoluble fraction, and refold it (see Material and Methods section for details).

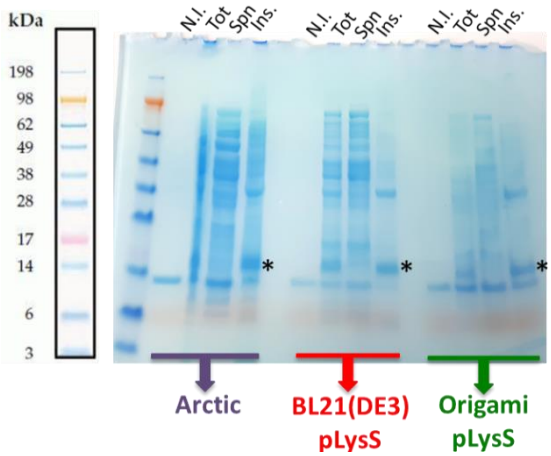


Figure 3.4. SDS-PAGE analysis of HT-Pvfp-5β expression in different cell strains. **N.I.:** no induced sample. **Tot:** Induced sample. **Spn:** soluble fraction after cells lysis. **Ins.:** inclusion bodies fraction after cells lysis. The apparent molecular mass of HT-Pvfp-5 on SDS-PAGE was 13 kDa versus the predicted molecular mass of 9.5 kDa. The difference is likely the consequence of the basic pI value (9.2) and a possible nonglobular shape of the protein.

The final optimized protocol for protein extraction and refolding, consists in extracting the protein in presence of 8M urea and reducing conditions, collecting it through a single step of IMAC affinity purification, and allowing correct refolding of the protein by gradual removal of urea and reducing agent in the presence of both reduced and oxidized glutathione (GSH) (Fig. 3.5).

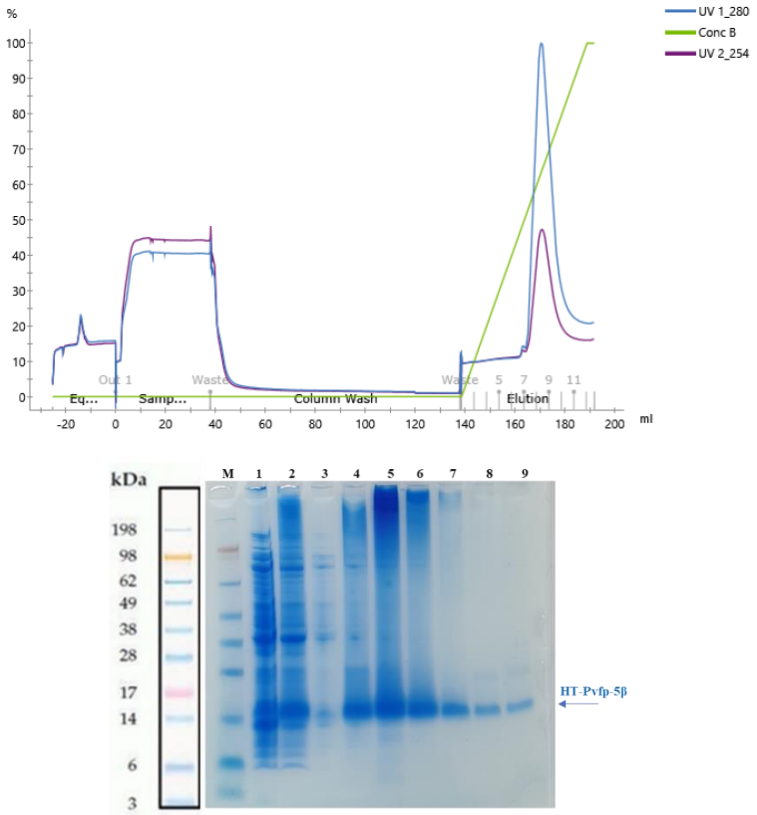


Figure 3.5. IMAC purification of HT-Pvfp-5β in detaturing and reducing conditions. **Top.** IMAC chromatogram. **Bottom.** SDS-PAGE analysis of IMAC fractions: (M) Marker; (1) Total; (2) Total extract; (3) IMAC fraction 6; (4) IMAC fraction 7; (5) IMAC fraction 8; (6) IMAC fraction 9; (7) IMAC fraction 10; (8) IMAC fraction 11; (9) IMAC fraction 12. HT-Pvfp-5β elutes at 65% of buffer containing 500 mM Imidazole.

His-tag removal by TEV protease digestion was performed after refolding, and it was achieved in phosphate buffer at pH 7.4 or sodium acetate buffer at pH 5.6 (Fig. 3.6, lanes 2 and 3). For both HT-Pvfp5β and cleaved Pvfp5β, the apparent molecular mass on SDS-PAGE was higher than expected. The difference is likely the consequence of the basic pI value (9.2) and a possible non globular shape of the protein. Cleaved protein was purified by reverse IMAC chromatography, but due to its adhesive properties Pvfp5β did not come

out into the column flow through as expected, but it eluted at 15% of elution buffer as pure protein, and at 20% of elution buffer together with the cleaved tag (Fig. 3.7).

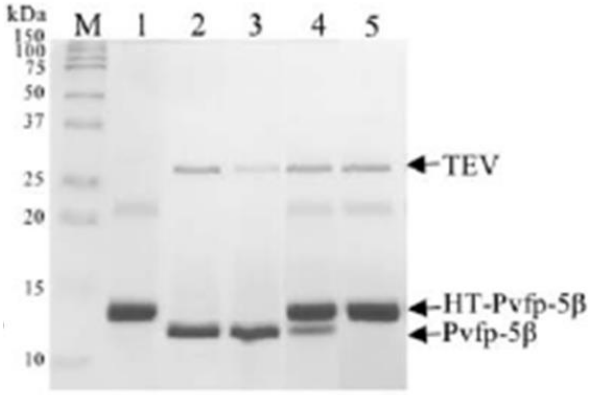


Figure 3.6. SDS-PAGE analysis of TEV cleavage test on HT-Pvfp-5β. (M) protein marker; (1) HT-Pvfp-5β in PB pH 7.4; (2) HT-Pvfp-5β in PB pH 7.4 plus TEV; (3) HT-Pvfp-5β in 0.1 M sodium acetate buffer pH 5.6 plus TEV; (4) HT-Pvfp-5β in 0.1 M sodium acetate buffer pH 4 plus TEV; (5) HT-Pvfp-5β in 5% acetic acid pH 2.0 plus TEV. Image published on Santonocito *et al.* (2019)[55]

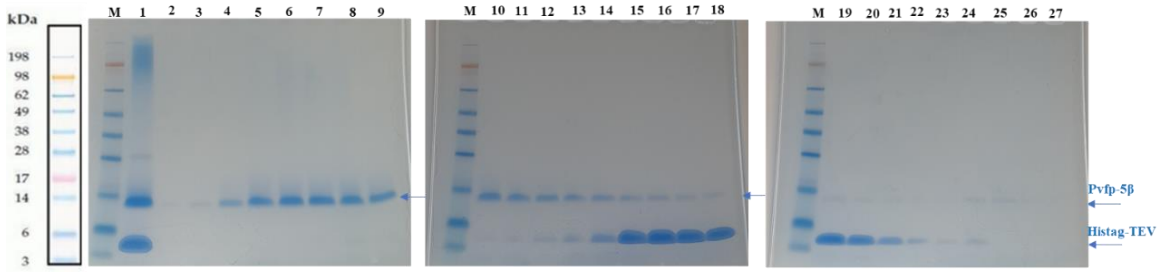


Figure 3.7. SDS-PAGE analysis of reverse IMAC fractions performed to recover Pvfp-5β after TEV cleavage. (M) Marker; (1) Cleaved Pvfp-5β before IMAC; (2) IMAC wash; (3-12) IMAC fractions from 2 to 11 eluted at 15% Elution Buffer; (13-18) IMAC fractions from 12 to 17 eluted at 20% Elution Buffer; (19-23) IMAC fractions from 18 to 22 eluted with 20% Elution Buffer; (24-27) IMAC fractions from 23 to 26 eluted with 100% Elution Buffer.

3.3 Probing fold of recombinant HT-Pvfp-5β by CD and NMR

The fold of recombinant Pvfp-5β was investigated in the range of pH 2.0-5.6 using his-tagged protein. Far-UV CD spectrum of the protein dissolved in 5% acetic acid (pH 2.0) exhibited a negative peak at 200 nm and a weak positive maximum at 230 nm, the latter probably caused by the high content of tyrosines (17.5%) (Fig. 3.8 A). The spectra did

not change at less acidic pH, except for a minor increase of the maximum at 230 nm and a small deepening of the minimum at 200 nm. Notably, CD spectrum of recombinant HT-Pvfp-5β is very similar to that published for wild type protein directly purified from mussels[54], confirming that the recombinant protein has properties similar to those observed for the protein from natural sources. However, the CD spectra resulted far from what is expected for a β-sheet protein[85], these results were considered inconclusive for deciding whether the protein has the EGF-like fold expected. A ^{1}H two-dimensional NOESY NMR spectrum was then recorded, and it revealed the presence of a well distribution of amide protons, which resonate up to 9.6 ppm. Moreover, NOESY spectrum revealed the presence of spatial nuclei correlation, which are close in space, and are generally related to a β-rich protein such as EGF-like domains (Fig. 3.8 B).

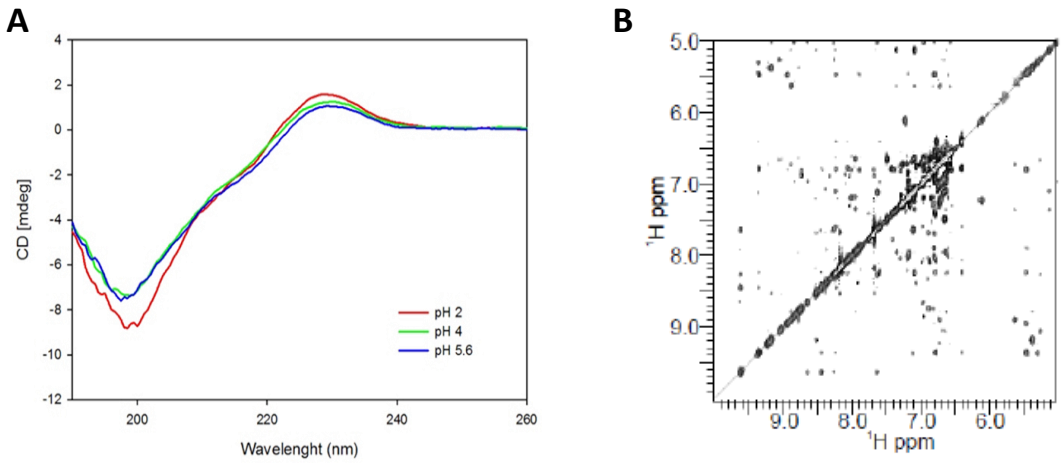


Figure 3.8. Probing fold of HT-Pvfp-5β. **A.** CD spectrum of the protein in 5% acetic acid (pH 2.0, red line), 0.1 M sodium acetate buffer (pH 4.0, green line) and 0.1 M sodium acetate buffer (pH 5.6, blue line). The final concentration of the tested protein was 34 µM. All spectra were corrected by solvent subtraction. **B.** Portion of the homo-nuclear 2D NOESY spectrum. Images published on Santonocito *et al.* (2019)[55]

3.4 Identification of the disulfide patterns in HT-Pvfp-5β by MS

A MS-based approach was used to characterize the disulfide bonds pattern of HT-Pvfp-5β. This work was carried out in collaboration with Prof Fabrizio Dal Piaz of University of Salerno. To ascertain which cysteines were involved in disulfide bonds, HT-Pvfp-5β was treated with iodoacetamide and then analyzed with ESI-MS. This indicated that only 2 of the 12 cysteines are not involved in disulfide bridges. The carboxyamidomethylated

HT-Pvfp-5β was then treated with digestion enzymes (tripsin and chymotrypsin), and the mixture was analyzed by LC-MS. Analysis of the resulting peptides, allowed description of the coupling of the 10 cysteines involved in covalent bonds (Table 3.1). In particular, the following disulfide bridges were found: Cys13-Cys29, Cys31–Cys40, Cys45–Cys56, Cys50–Cys67, and Cys69–Cys78. Cys8 and Cys19 resulted not involved in a disulfide bridge and were found as carboxyamidomethylated peptides-(4–11) and (15–20), respectively. These results are consistent with what it was obtained from the bioinformatic studies described in paragraph 3.1, and confirm a high homology with two tandem EGF-like motifs. However, in contrast with the structural model obtained by SWISSMODEL, mass spectrometry analysis revealed the presence of the disulfide bridge between Cys45–Cys56 suggesting that the accuracy of the homology modeling in this region could be low (Fig. 3.9).

Peptide	Experimental molecular weight	Theoretical molecular weight	Involved cysteines
4-11	996,409	996,405	Cys8-CAM*
(12-23)-(29-30)	656,281	656,255	Cys13-Cys29
15-20	635,282	635,270	Cys19-CAM*
(31-32)-(39-42)	801,333	801,305	Cys31-Cys40
(43-46)-(52-58)	1121,469	1121,404	Cy45-Cys56
(47-51)-(67-68)	568,326	568,303	Cys50-Cys67
(69-71)-(77-80)	869,374	869,356	Cys69-Cys78

Table 3.1. Mass spectrometry-based identification of cysteine-containing peptides generated by tripsin and chymotrypsin digestion of carboxyamidomethylated HT-Pvfp-5β. *Carboxyamidomethylated cysteine. Data published on Santonocito *et al.* (2019)[55]

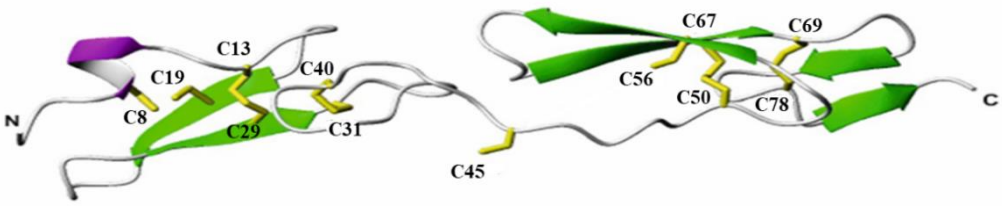


Figure 3.9. Structural model of HT-Pvfp-5β based on sequence similarity with Notch ligand delta-like 1 protein. The disulfide bridge Cys45–Cys56 is not formed in the model as the consequence of a two-residue deletion with respect Notch sequence, but it is experimentally found by mass spectrometry analysis. Image published on Santonocito *et al.* (2019)[55]

3.5 Cell adhesion, spreading abilities and citotoxity effect of HT-Pvfp-5β

3.5.1 Surface coating of HT-Pvfp-5β

In order to test adhesive proprieties of HT-Pvfp-5β, its surface coating ability was investigated in collaboration of Dr Rosa Passantino of IBF-CNR, Palermo. The method used is based on the observation that *mfps* come out of solution as pH is raised, and spontaneously adsorb and coat to the first surface they contact. Hydrophilic surface, hydrophobic surface, and glass surface were tested. The analysis was performed using Cell-Tak, that is a natural extract of *Mytius edulis* adhesive-proteins, and uncoated wells as positive and negative controls, respectively. HT-Pvfp-5β showed remarkable coating ability for all surfaces, and a more pronounced staining, than with Cell-Tak, was observed on both polystyrene surfaces (Fig. 3.10).

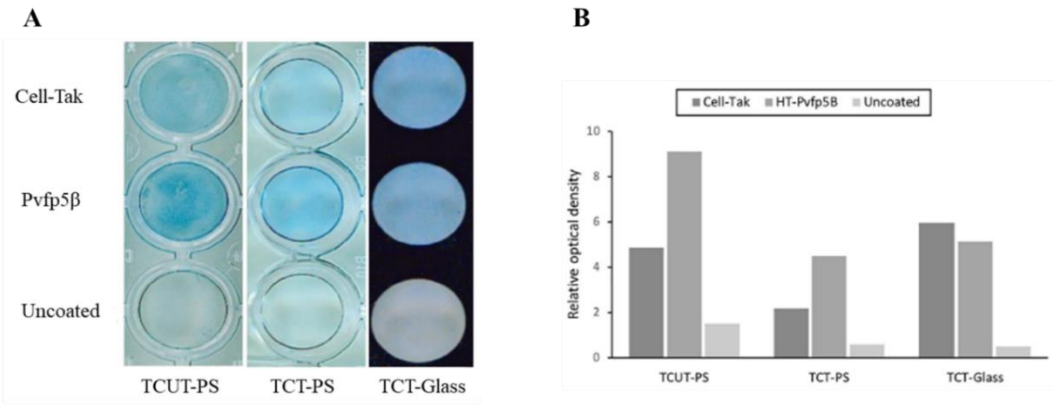


Figure 3.10. Coating analysis of HT-Pvfp-5β on Polystyrene and Glass surfaces. Cell-Tak and uncoated wells were used as positive and negative controls, respectively. The figure shows the Coomassie blue staining (A) and the relative densitometry analysis (B) of the coated proteins. Image published on Santonocito *et al.* (2019)[55].

3.5.2 Effect of HT-Pvfp-5β on cells viability, adhesion and spreading

The surface coating ability of HT-Pvfp-5β led to investigate the effect of HT-Pvfp-5β coating on cells viability, adhesion and spreading. This was done in collaboration with Dr Rosa Passantino of IBF-CNR, Palermo.

The effect of HT-Pvfp-5β on cells viability was assessed by MTS assay. This is a colorimetric method for sensitive quantification of viable cells in proliferation and cytotoxicity, and it is based on the reduction of the MTS tetrazolium compound by viable

cells[145]. This reaction generates formation of a coloured formazan product in cell culture media that is induced by NAD(P)H-dependent dehydrogenase enzymes. Absorbance of the formazan dye is then evaluated at 490 nm wavelength, and its value is directly proportional to the number of viable cells in the culture. The effect of HT-Pvfp-5β on cells viability was assessed for NIH-3T3 and HeLa cells. Cells were seeded into uncoated or HT-Pvfp-5β coated wells of TCT-PS 96-well plates, and viability was tested up to 72 h. Both NIH-3T3 and HeLa cells remained viable after being into contact with HT-Pvfp-5β, and no statistically significant differences were observed between uncoated or HT-Pvfp-5β coated wells, indicating no cytotoxic effects of HT-Pvfp-5β (Fig. 3.11).

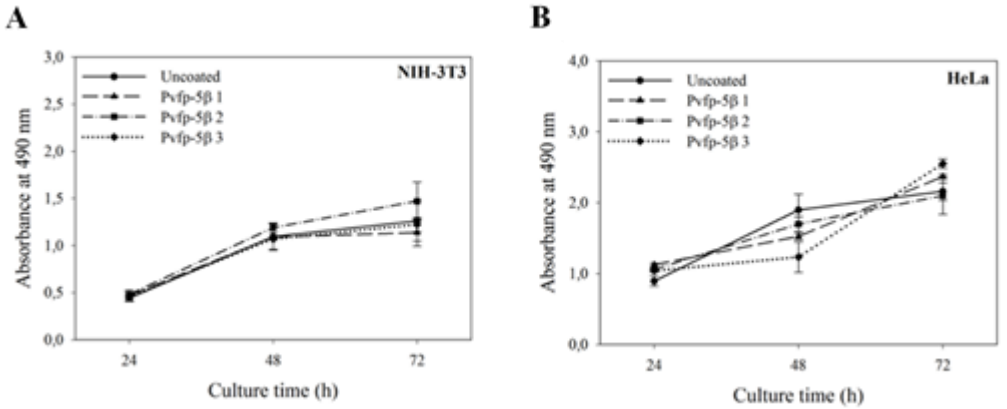


Figure 3.11. Cell viability of NIH-3T3 (A) and HeLa (B) cells grown on HT-Pvfp-5β coating wells. Three coating concentrations were used for HT-Pvfp-5β (1.75 µg/cm2 Pvfp-5β 1; 3.5 µg/cm2 Pvfp-5β 2; 7 µg/cm2 Pvfp-5β 3). Uncoated surface was used as negative control. NIH-3T3 and HeLa cells in serum-containing medium were seeded at a density of 5x103/well and incubated for 72 h. Each value represents the mean of three independent experiments ±the standard deviation. Image published on Santonocito *et al.* (2019)[55].

The effect of HT-Pvfp-5β on cells adhesion was assessed testing three different surfaces coated with HT-Pvfp-5β (glass and polystyrene). Wells coated with Cell-Tak and PLL were used as positive control, while non-coating wells were used as negative control. HeLa and NIH-3T3 cells were seeded into uncoated and coated wells of TCT-PS, TCUT-PS and TCT-G 96-well plates. After 2 h of cell-adhesion, the unattached cells were aspirated gently and rinsed with PBS. The MTS assay was performed to measure the living-cell adhesion quantitatively. HT-Pvfp-5β exhibited better cell-adhesion ability than PLL and Cell-Tak on glass and polystyrene plate with both cell lines. The number of adherent NIH-3T3 cells increased by about 4.2-6 folds on HT-Pvfp-5β coated glass

surface compared to the uncoated glass (Fig. 3.12A) and by 1.5-2.3 folds on HT-Pvfp-5β coated polystyrene surfaces (Fig. 3.12 B and C). Although HeLa cells appeared more spread on all surfaces coated with HT-Pvfp-5β, the number of adherent cells was higher only on HT-Pvfp-5β coated tissue-culture untreated polystyrene surfaces (1.8-1.9 folds) (Fig. 3.12 B).

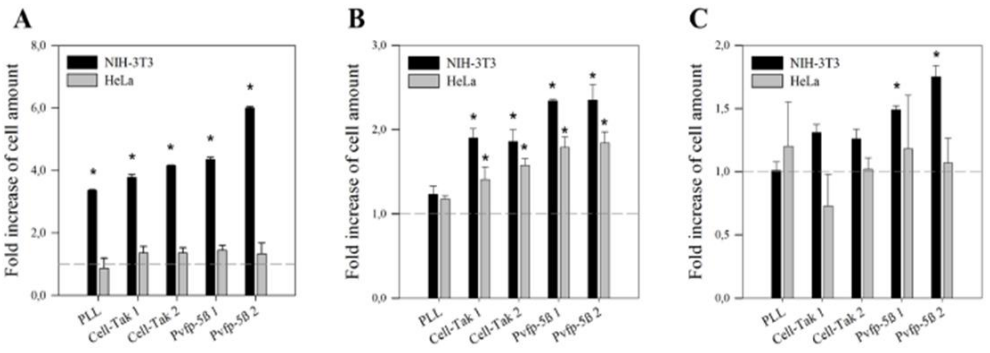


Figure 3.12. Cell adhesion of NIH-3T3 and HeLa cells on different surface materials coated with PLL, Cell-Tak or HT-Pvfp-5β. (A) tissue-culture treated glass plate TCT-G, (B) tissue-culture untreated polystyrene plate TCUT-PS, (C) tissue-culture treated polystyrene plate TCT-PS. Two coating concentrations were used for HT-Pvfp-5β and Cell-Tak (3.5 µg/cm^2 Pvfp-5β 1 and Cell-Tak 1; 7 µg/cm^2 Pvfp-5β 2 and Cell-Tak 2), PLL was used as positive control at 7 µg/cm^2. NIH-3T3 and HeLa cells in serum-free medium were seeded at a density of 5x10^4/well and incubated for 2 h. Fold increase of cell amount is referred to the uncoated control for each surface material (dotted gray line). The effects were determined by the MTS assay. Each value represents the mean of three independent experiments ±SD. Statistical significance: *P < 0.05 vs. uncoated control. Image published on Santonocito *et al.* (2019)[55]

After cell-attachment on surface, cells underwent morphologic changes and spreading with an active reorganization of the cytoskeleton. This is due to actin polymerization from a monomeric G-actin form to the filamentous polymeric F-actin form which causes rearrangements characterized by formation of protrusion including filopodia and lamellipodia. To study whether the adhesive substrate HT-Pvfp-5β was capable of supporting spreading, F-actin was stained to observe the cytoskeleton. HeLa and NIH-3T3 cells were seeded into uncoated, PLL or HT-Pvfp-5β coated glass wells. Unattached cells were aspirated and the adhesion state and spreading of attached cells were analyzed by actin staining by the green-fluorescent Alexa Fluor™ 488 Phalloidin and fluorescence microscopy. A higher quantity of adherent cells with a more organized cytoskeleton were observed on HT-Pvfp-5β coating in both HeLa and NIH-3T3 cells if compared to the poorly spread cells on PLL coated and uncoated wells (Fig. 3.13). These results show that coating of HT-Pvfp-5β is biocompatible and significantly affects cell attachment and spreading.

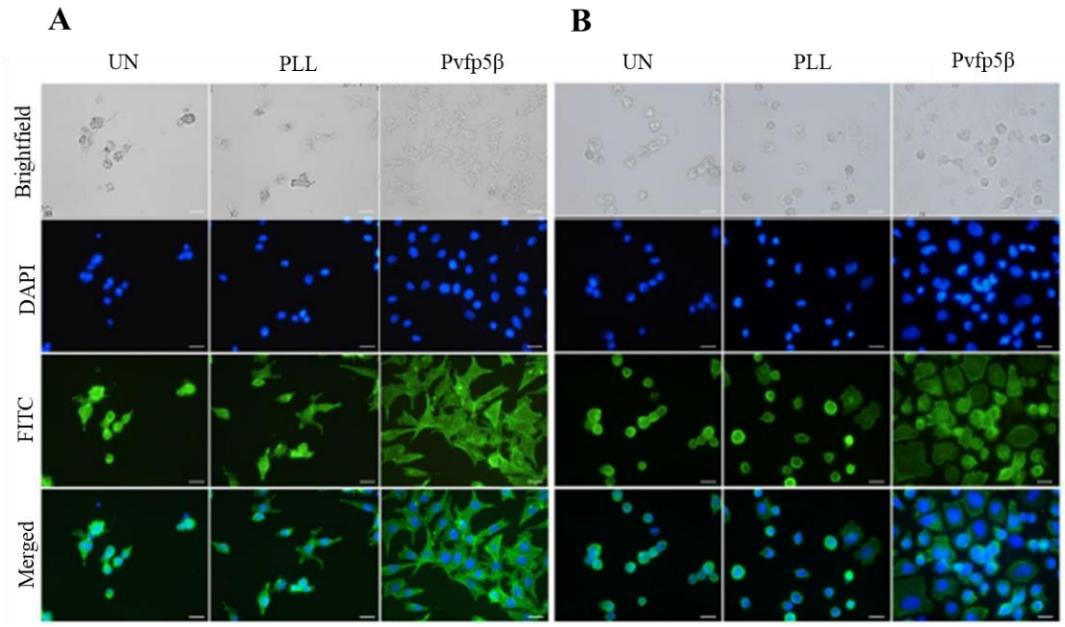


Figure 3.13. Cell spreading of NIH-3T3 and HeLa cells on uncoated (UN), Poly-L-Lysine (PLL) or HT-Pvfp5b coated glass well. Fluorescent microscopy images of NIH-3T3 cells (A) and HeLa cells (B). Cell nuclei was in blue (Hoechst staining-DAPI channel) and F-actin was in green (Alexa Fluor™ 488 labelled phalloidin staining-FITC channel). Scale bar: 20 μm. Image published on Santonocito *et al.* (2019)[55].

3.6 NMR structure determination for Pvfp-5β

In this research work, the structure of Pvfp-5β was determined by NMR. Indeed, Pvfp-5β is a quite small protein (9.5 KDa), and it is then suited for NMR studies. This was a great advantage because it allowed to perform analysis in solution thus achieving structural information closer to the native situation. Moreover, the choice of NMR in this research work allowed to perform dynamic studies in solution of Pvfp-5β. All NMR measurements were acquired on ^{15}N,^{13}C-labelled protein samples dissolved in acetate buffer at pH 4.5 such to maintain Pvfp-5β in its monomeric form. Several 2D ^{15}N,^{1}H HSQC spectra were acquired at different temperatures and times, and showed that at pH 4.5 Pvfp-5β remains well folded and stable for months and has no tendency to aggregate in the tested conditions. In order to prove that cellular assays were not altered by the presence of His-tag, ^{15}N,^{1}H HSQC spectra were also recorded both on HT-Pvfp-5β and cleaved Pvfp-5β. They resulted well overlapped confirming that Pvfp-5β folding is not affected by the His-tag (Fig. 3.14).

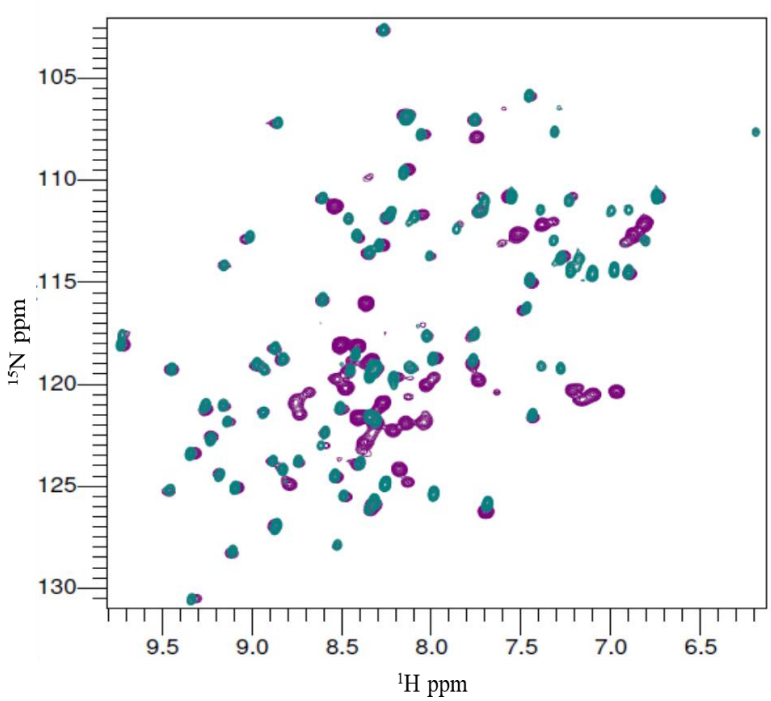


Figure 3.14. Overlap of ^{15}N,^{1}H HSQC spectra of HT-Pvfp-5β (purple) cleaved Pvfp-5β (teal). The extra peaks in purple are related to the His-tag residues.

3.6.1 Backbone resonances assignment

^{15}N,^{1}H HSQC, HNCACB, HN(CO)CACB, HN(CO)CA, HNCA, HNCO, HBHA(CO)NH, and HBHANH spectra were recorded, processed and analyzed to assign Pvfp-5β backbone resonances. In particular, ^{15}N,^{1}H HSQC peaks allowed to identify the resonances of proton and nitrogen amide backbone groups. These resonances were then assigned to the corresponding amino acids by analysis of HN(CO)CACB, HN(CO)CA and HNCA, HNCACB experiments, which also provided chemical shift values of Cα and Cβ atoms. Indeed, there are residues very easy to spot, like serine, threonine, alanine and glycine as their Cα and Cβ chemical shifts are very different to those of other amino acids, and in the case of glycine there is no Cβ. The identification of these amino acid types allows the sequence-specific assignment (Fig. 3.15). 73 out of 75 non-proline residues were assigned and the corresponding assigned ^{15}N,^{1}H HSQC spectrum is reported in Fig. 3.16. The HNCO spectrum was used to assign the carbonyl carbon chemical shifts, while chemical shift values of Hα and Hβ were assigned by analysis of HBHA(CO)NH and HBHANH experiments (Fig. 3.17). Full backbone assignment is reported in ANNEX III.3 together with side chains assignment.

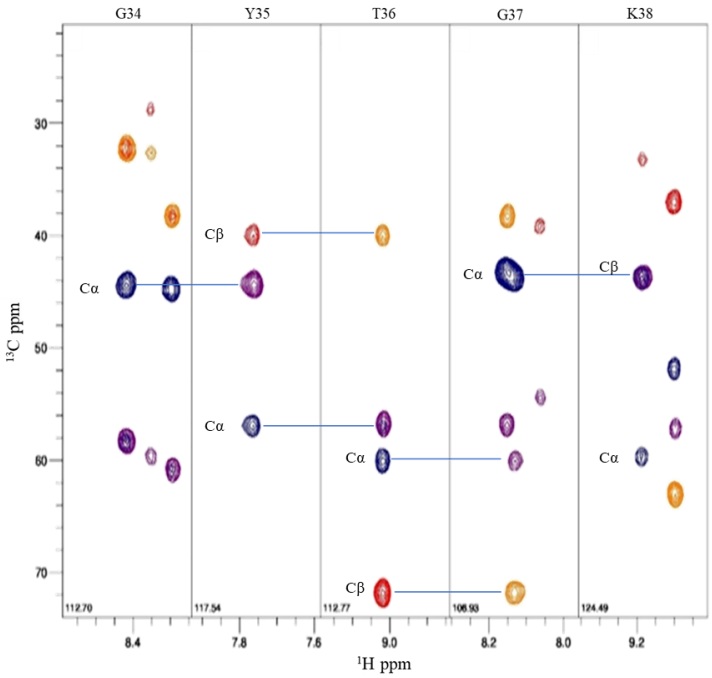


Figure 3.15. Strips showing sequential connectivity for residues G34 to K38. Strips from CBCA(CO)NH (positive peaks in purple. Negative peaks in orange) are overlapped with those from HNCACB (positive peaks in blue. Negative peaks in red).

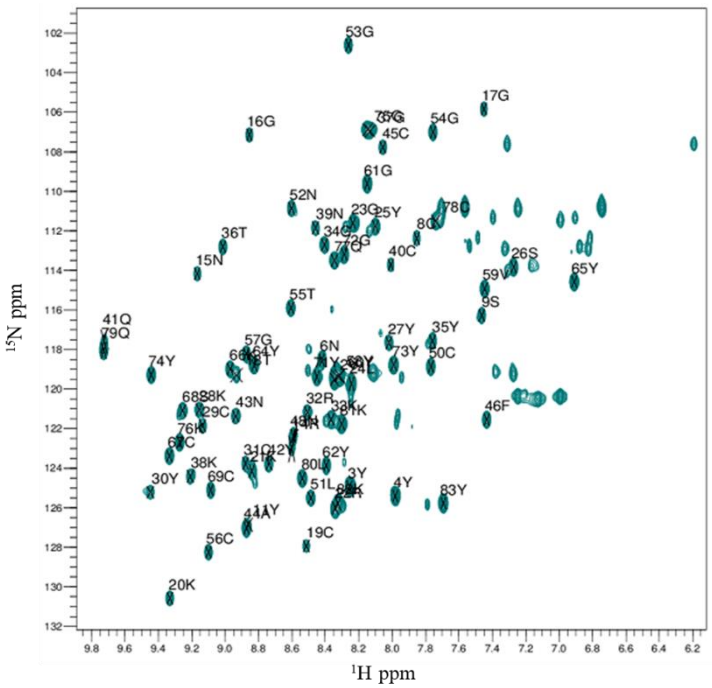


Figure 3.16. ^{15}N,^{1}H HSQC spectrum of recombinant Pvfp-5β. 97% of the backbone resonances has been assigned. Peaks are labelled with single amino acid letter code and corresponding residue number.

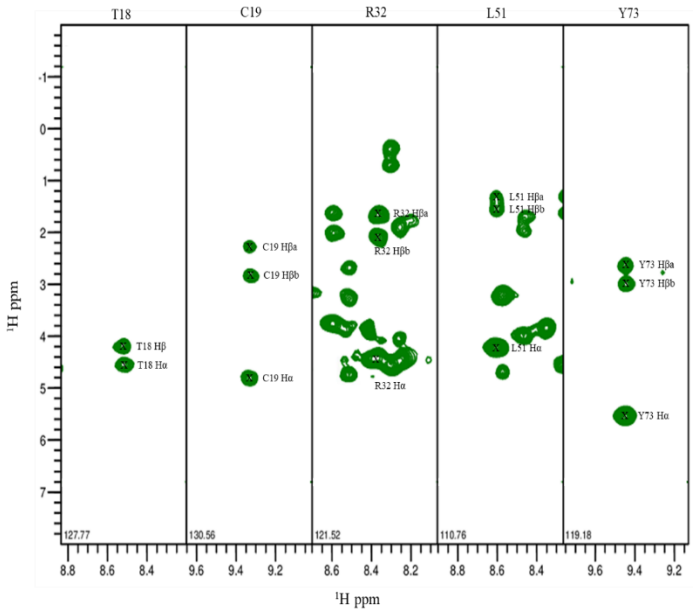


Figure 3.17. Strips from HBHA(CO)NH spectrum, illustrating Hα and Hβ peaks for T18, C19, R32, L51 and Y73 residues.

3.6.2 Secondary structure prediction

Backbone chemical shift values of HN, HA, CA, CB, CO and N were used in TALOS+ software for empirical prediction of dihedral angles ψ and ϕ, and secondary structure. The secondary structure predicted by TALOS+ is consistent with what previously obtained from the homology structural model, showing the presence of β-sheets and random coil structure in Pvfp-5β (Fig. 3.18).

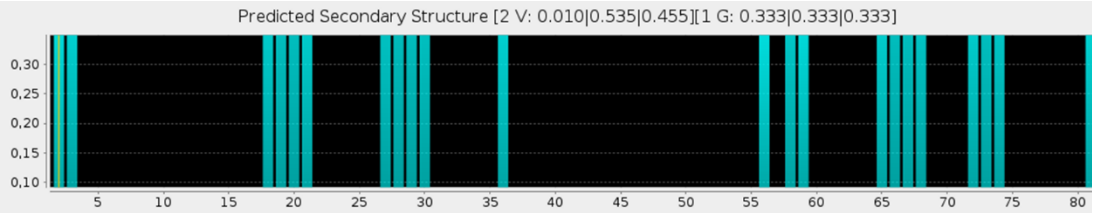


Figure 3.18. Secondary structure prediction by TALOS+. Blue bars indicate those residues involved in β-strand structures. All the other residues are in random coil structure.

The predicted Random Coil Index (RCI) S^2 values (Fig. 3.19), reveal the presence of flexible N-terminal and C-terminal regions, while mainly protein is more rigid, as expected for a protein with β-sheets structure stabilized by five disulfide bridges.

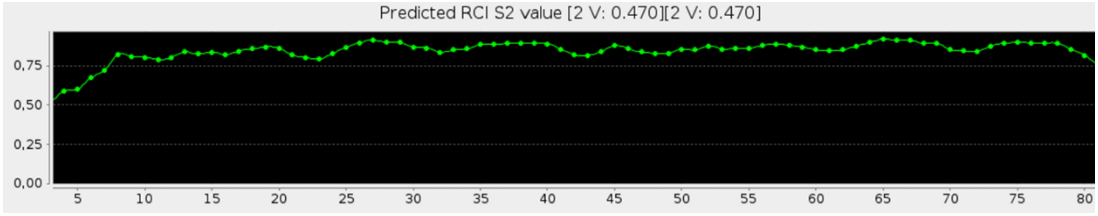


Figure 3.19. Order parameter S^2 values calculated from chemical shifts by TALOS+.

The quality of predicted dihedral angles for Pvfp-5β was analysed by Ramachandran plot[146], where they are reported energetically allowed regions for backbone dihedral angles ψ against φ of each amino acid residue in the protein. The Ramachandran plot of the 10 best database matches given for Pvfp-5β residues involved in β-strand structure (T18, K28, Y58 and Y65) is shown in Fig. 3.20.

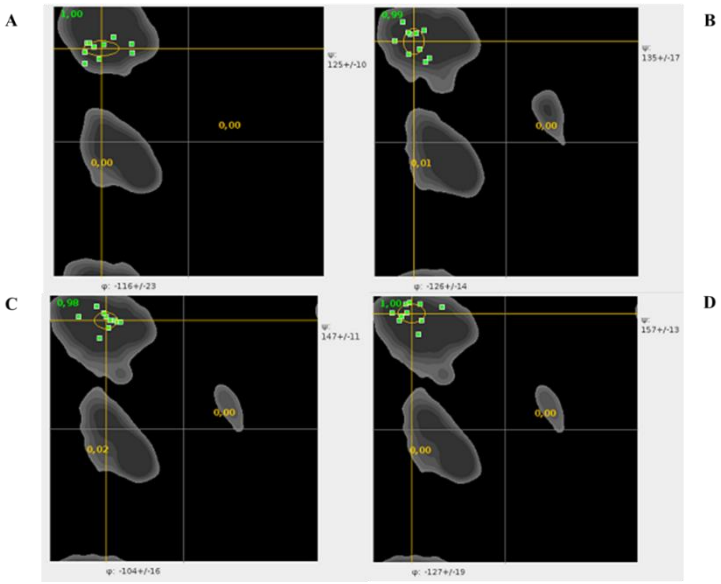


Figure 3.20. Ramachandran plot by TALOS+ showing ψ/φ distributions of the 10 best database matches for Thr 18 (A), Lys 28 (B), Tyr 58 (C), and Tyr 65 (D).

3.6.3 Side chains resonances assignment

The side-chain assignment was performed by analysis of 3D HCCH-TOCSY experiments, recorded both for aliphatic, and aromatic chains. 2D (HB)CB(CGCD)HD and (HB)CB(CGCD)CEHE spectra were used to assign Hδ and Hε protons of aromatic residues. In the HCCH-TOCSY, Cα and Cβ chemical shifts obtained from the backbone

assignment were used to identify the resonances of all the side chain carbons and related attached hydrogens (Fig. 3.21). Except for residue Proline 12, almost all the side chains of Pvfp-5β amino acid residues were successfully assigned. Conversely to aliphatic chains assignment, the assignment of aromatic side chains was quite difficult using both aromatic 3D HCCH-TOCSY, and 2D (HB)CB(CGCD)HD and (HB)CB(CGCD)CEHE experiments. In these cases, missing resonances were assigned by analysis of 3D NOESY spectra, such as ^{13}C NOESY-HSQC . Full side chains assignment is reported in ANNEX III.3 together with backbone assignment.

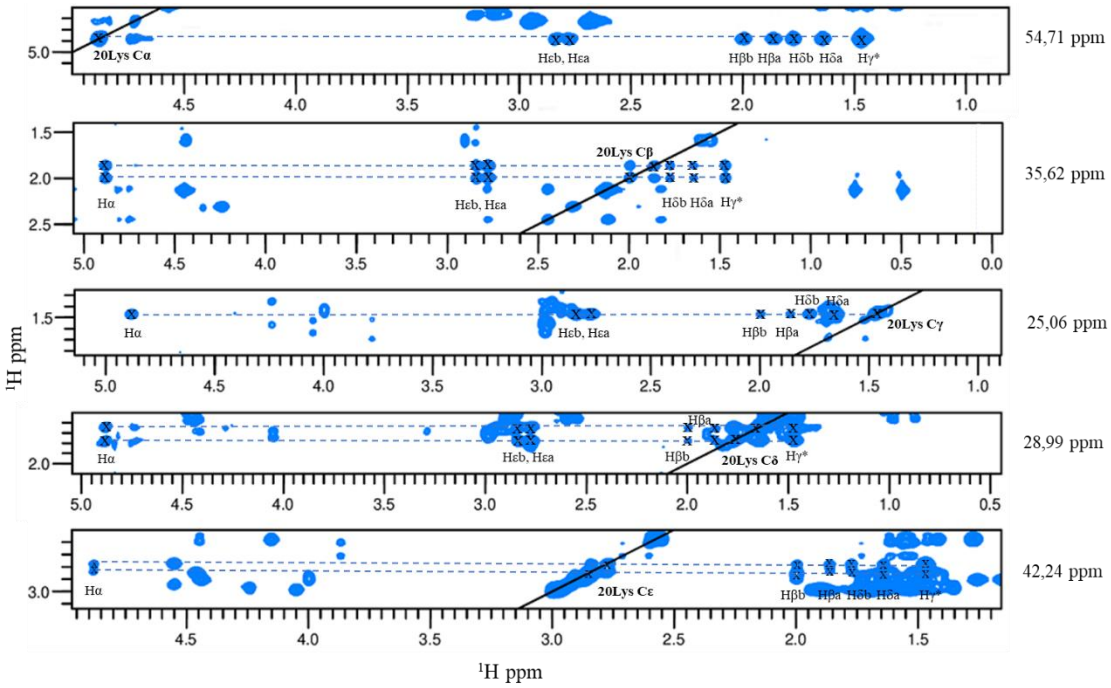


Figure 3.21. Strips from HCCH-TOCSY spectrum for Lys 20 side chains assignment.

3.6.4 NMR structure calculation

The three-dimensional structure of Pvfp-5β was calculated with ARIA 2.3[147] after eight iterations, with 50 conformers for iteration, and a final refinement step in water, where 20 conformers, with lowest energies, were obtained. Input data for ARIA 2.3 were Pvfp-5β amino acid sequence, chemical shift assignment list, NOE restraints lists, dihedral angles, disulfide bridges and hydrogen bonds. Disulfide bridges costraints were those identified by the mass spectrometry analysis: C13-C29, C31-C40, C45-C56, C50-C67,

C69-C78. Direct detection of N-H···O=C hydrogen bonds was attempted by acquisition and analysis of long-range HNCO spectrum, but they were tough to quantify due to the low sensitivity of this type of experiment. Hydrogen bond restraints used for Pvfp-5β structure calculation, were then chosen basing on β-sheets predicted by TALOS+: T18-Y30, K20-K28, T55-S68 and G57-Y66. NOE restraints were assigned manually by analysis of 3D ^{15}N NOESY-HSQC, 3D ^{13}C NOESY HSQC centred on aliphatic region, and 3D ^{13}C NOESY HSQC centred on aromatic region. Ambigous NOE assignments were sorted automatically by ARIA 2.3. ARIA was set to perform eight iterations, with 50 conformers for each iteration. A different violation threshold was set for each iteration: 200.0 for it0, 6.0 for it1, 3.0 for it2, 2.0 for it3, 1.0 for it4 and it5, 0.5 from it6 to it8. The conformers with the lowest energy values were analysed to filter the set of distance restraints for false positive and assign ambiguities. ARIA calculations were also performed with the adaptive choice of violation tolerance in order to adjust the violation tolerance automatically to the quality of the experimental data. NOESY peak list frequency windows were chosen as 0.02 ppm for ^{1}H, and 0.4 ppm for both ^{15}N and ^{13}C nuclei. The log-harmonic distance restraint potential was also added to ARIA protocol. This potential derives from a Bayesian analysis showing that NOEs and the derived distances ideally follow the log-normal distribution[148,149]. The log-harmonic potential is applied during the second cooling stage of the molecular dynamics simulated annealing (MDSA) and during water refinement.

The structure ensemble of the twenty best conformations obtained for Pvfp-5β is shown in Fig. 3.22, while the conformer at lowest energy (-2429.97 Kcal mol^{-1}) is reported in Fig. 3.23. For the 20 structures with the lowest energy, the backbone root-mean-square deviation (RMSD) was calculated and resulted to be 1.56 ± 0.40 Å, which indicates that structures are quite well overlapped. Further analysis is underway to try do decrease RMSD further. Overall, Pvfp5-β structure determined by NMR, is consistent with mass spectrometry data and also confirms the high homology with tow tandem EGF motifs. Pvfp-5β structure results mainly dominated of random coil, and includes two coupled antiparallel β-sheets held together by the presence of five disulfide bridges: the first β-sheet involves regions 17-21 and 27-30, while the second β-sheet involves regions 55-57 and 66-68. Eight violations are still occuring above to 0.5 Å, along with a low overlap affecting the N-teminal side, and traits 17-21 and 27-30 involved in β-sheet structure.

This could be explained by the presence of local motions occuring at ps-ns timescale. However, ispection of upper bound violations above 0.5 Å is still under refinement.

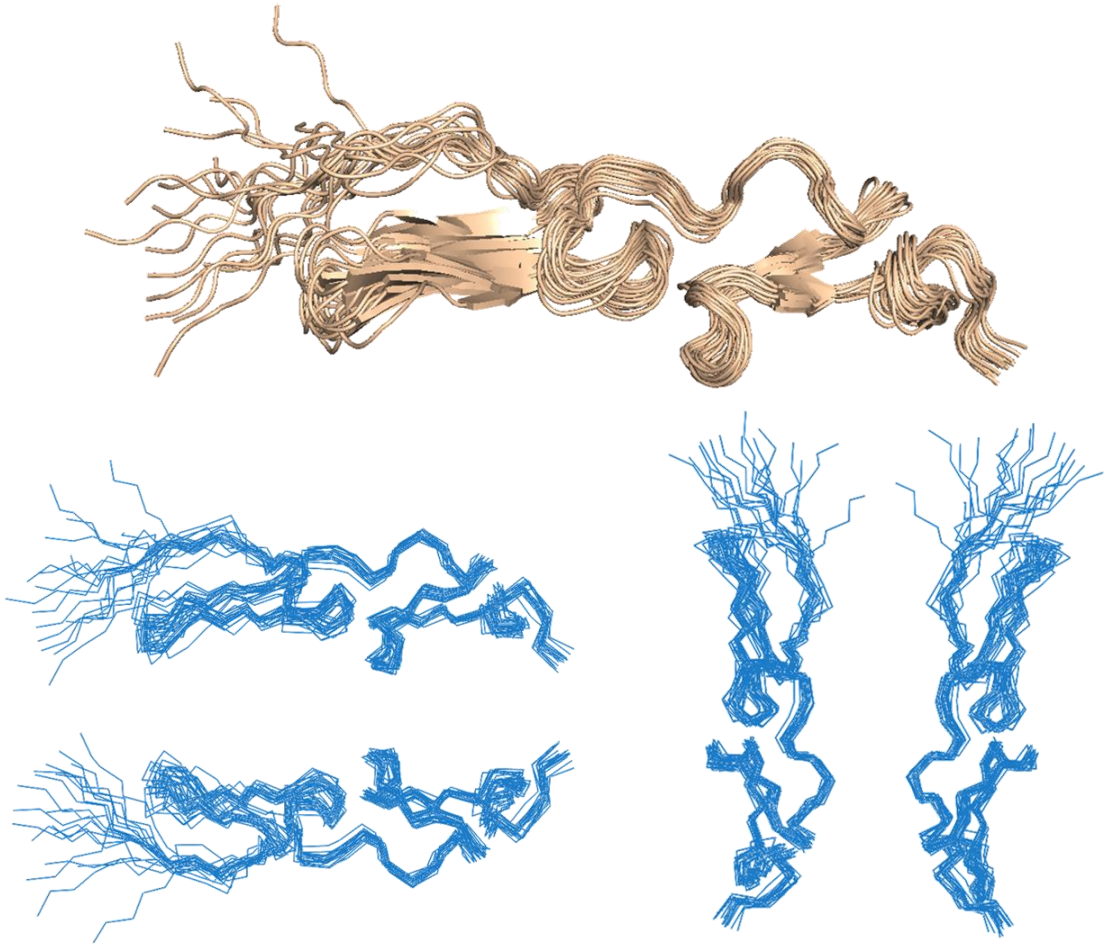


Figure 3.22. Conformation ensemble of Pvfp5-β, in cartoon (wheat) and ribbon (blue), determined by ARIA 2.3 software after water refinement.

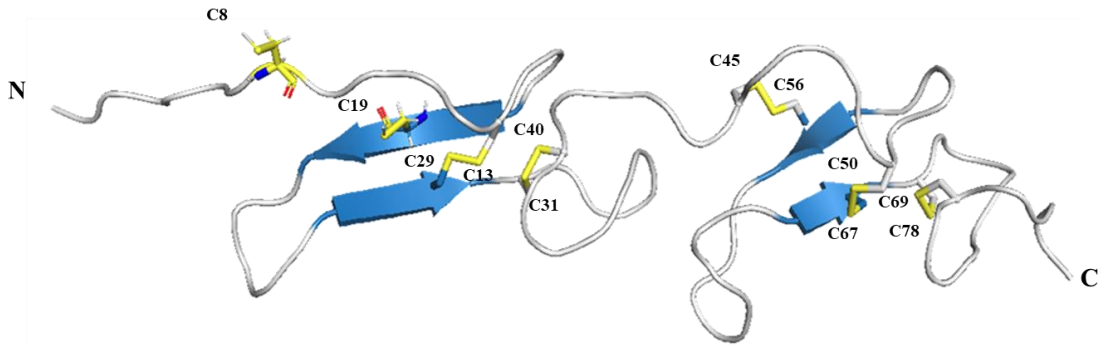


Figure 3.23. Three-dimensional lowest energy structure of Pvfp5-β showing the five disulfide bridges: C13-C29, C31-C40, C45-56, C50-C67, C69-C78. The two cysteines 8 and 19 are far to be involved in disulfide bridge.

The Pvfp-5β structure could be explained how this protein performs its interaction with marine surfaces. The protein is rich in basic amino acids, like lysines and arginines, all close in sequence to tyrosine residues that result modified in Dopa in the protein extracted from the natural source (Fig. 3.24). The positively charged residues appear all exposed to protein surface in Pvfp-5β NMR structure, and therefore they can be easily involved in saline bridges with negatively charged groups of marine surfaces, improving Pvfp-5β adhesion properties. At basic pH value, mussel plaque solidification occurs, which is justified by covalent interaction with surface, and could be explained also with higher nucleophilic reactivity of lysine and arginine. Thanks to the proximity to arginine and lysine residues, both in sequence and space, tyrosine residues can get close to the marine surface too and they can also be involved in mussel adhesion and cohesion.

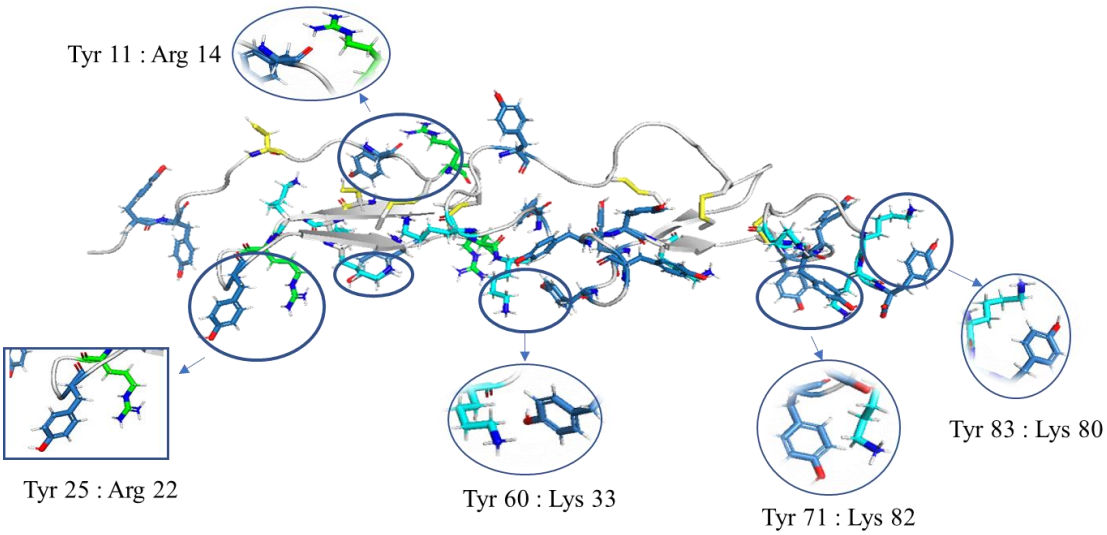


Figure 3.24. Three-dimensional structure of Pvfp5-β showing arginine and lysine residues close to tyrosine.

3.7 Molecular Dynamic studies on Pvfp-5β through NMR

To assess if Pvfp-5β is affected of internal motion, at ps-ns timescale, backbone ^{15}N relaxation measurements were performed at two static magnetic fields (600 and 800 MHz). In particular, to determine the global molecular tumbling and local correlation times, T1, T2 and HetNOE values were calculated (full T1, T2 and HeteroNOE analysis

is reported in ANNEX III.4). Although Pvfp-5β is largely composed of a random coil structure, the data indicate that, overall, the protein is well ordered, as revealed from uniform values of T1, T2, and HetNOE, at both 600 and 800 MHz (Fig. 3.25-3.30). This qualitative interpretation can lead to consider Pvfp-5β as a rigid protein with fluctuations in the order of ps-ns occur. HetNOE experiments provided information about the motion of individual NH bond vectors, and revealed that the following residues undergo motion faster than the overall tumbling of the protein: V2, Y3, Y4, N6, G23, L24, Y25, S26, S33, G61, and Y83. The ratio T1/T2 revealed, instead, the presence of residues that could be affected by fluctuations at µs-ms timescale: V2, Y3, Y4, N6, C8, T18, G23, L24, Y27, N39, Y42, C45, G61, and Y83 (Fig. 3.31-3.32). All these data justify the presence of disorder in the N-terminal region of Pvf-5β, the lack of a good overlap of the conformers' ensemble obtained by structure calculation, and the huge effort needed to refine Pvf-5β NMR structure.

The ratio T1/T2 of 37 residues, which showed fixed values of T1 and T2, was exploit to calculate the correlation time of Pvf-5β, using Eq.3.1, where ν_N is the frequency of nitrogen at the used magnetic field. The resulted value for correlation time was 7.5 ns at both 600, and 800 MHz, and it can justify for a monomeric and ellipsoidal protein with a molecular weight of 9.5 KDa, as is Pvfp-5β.

Eq. 3.1 $$\tau c \cong \frac{1}{4\pi\nu N}\sqrt{6\frac{T1}{T2}-7}$$

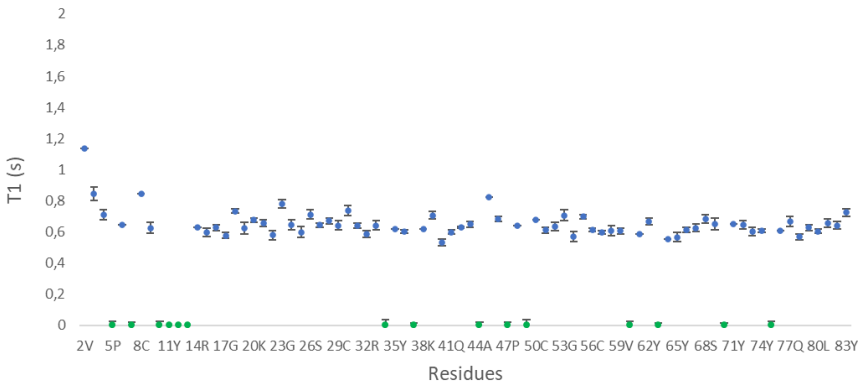


Figure 3.25 Longitudinal relaxation measured at 600 MHz. In green residues for which no T1 was calculated due to the presence of peaks overlap (exception is made for proline non observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 34G, 37G, 44A, 47P, 49P, 60Y, 63P, 70P and 78C.

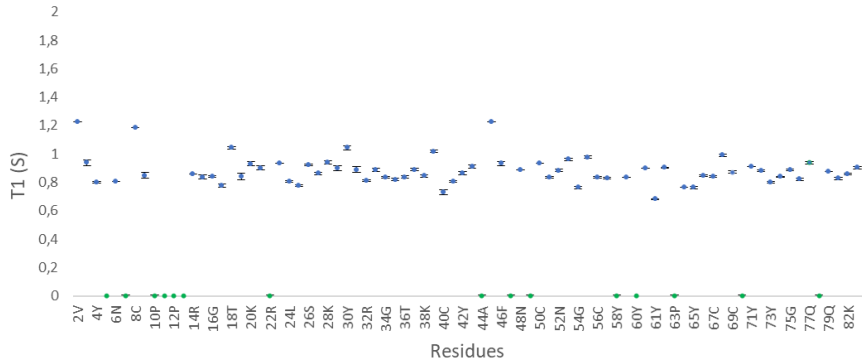


Figure 3.26 Longitudinal relaxation measured at 800 MHz. In green residues for which no T1 was calculated due to the presence of peaks overlap (exception is made for proline non observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 22R, 44A, 47P, 49P, 58Y, 60Y, 63P,70P and 78C.

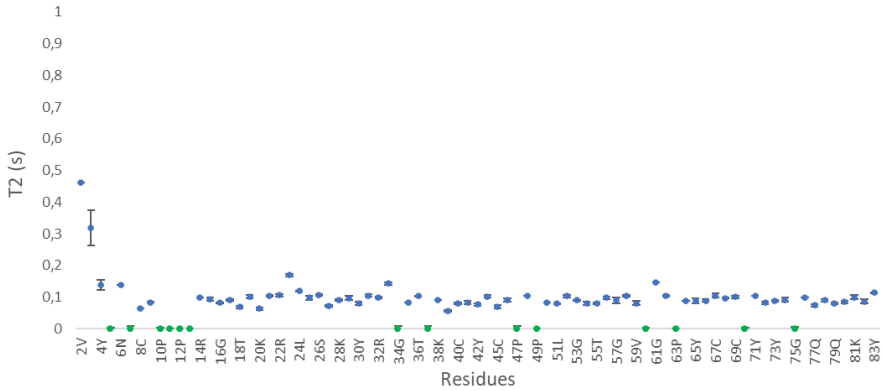


Figure 3.27 Transverse relaxation measured at 600 MHz. In green residues for which no T2 was calculated due to the presence of peaks overlap (exception is made for proline no observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 34G, 37G, 47P, 49P, 60Y, 63P, 70P and 75G.

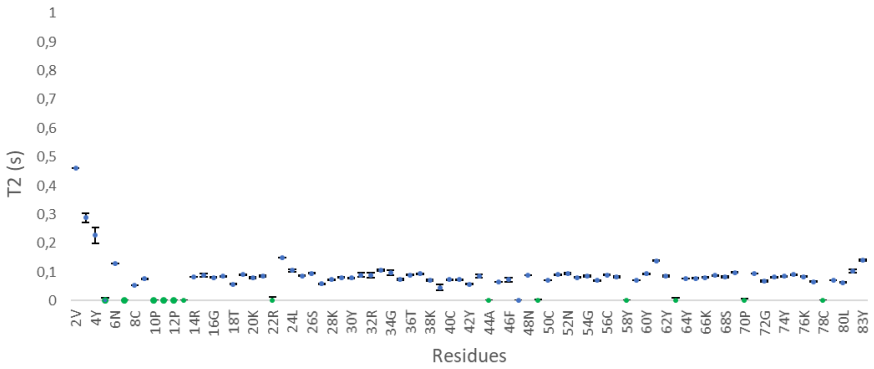


Figure 3.28 Transverse relaxation measured at 800 MHz. In green residues for which no T2 was calculated due to the presence of peaks overlap (exception is made for proline non observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 22R, 44A, 47P, 49P, 58Y, 63P, 70P and 78P.

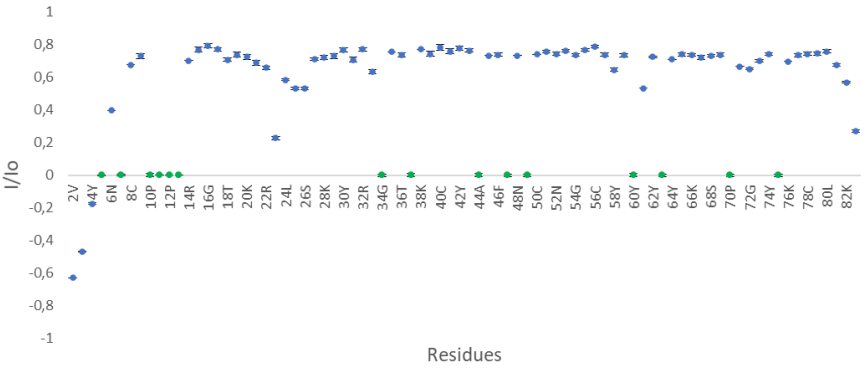


Figure 3.29 HeteroNOE measured at 600 MHz. In green residues for which no HeteroNOE was calculated due to the presence of peaks overlap (exception is made for proline no observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 34G, 37G, 44A, 47P, 49P, 60Y, 63P, 70P and 75G.

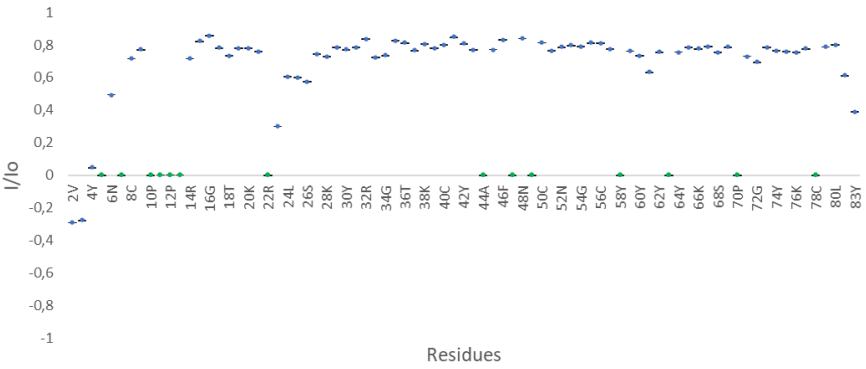


Figure 3.30 HeteroNOE measured at 800 MHz. In green residues for which no HeteroNOE was calculated due to the presence of peaks overlap (exception is made for proline no observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 22R, 44A, 47P, 58Y, 63P, 70P and 78C.

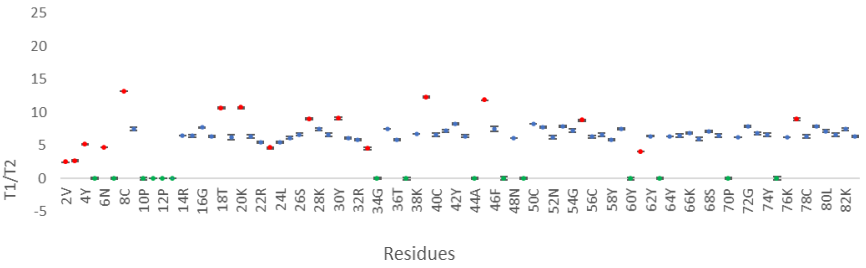


Figure 3.31 T1/T2 ratio measured at 600 MHz. In green residues for which no T1 and T2 values were calculated due to the presence of peaks overlap (exception is made for proline no observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 34G, 37G, 44A, 47P, 49P, 60Y, 63P, 70P and 75G. In red those residues whose T1/T2 ratio value is outside the observed range: 2V, 3Y, 4Y, 6N, 8C, 18T, 20K, 23G, 27Y, 30Y, 33Y, 39N, 45C, 55T, 61G, 77Q.

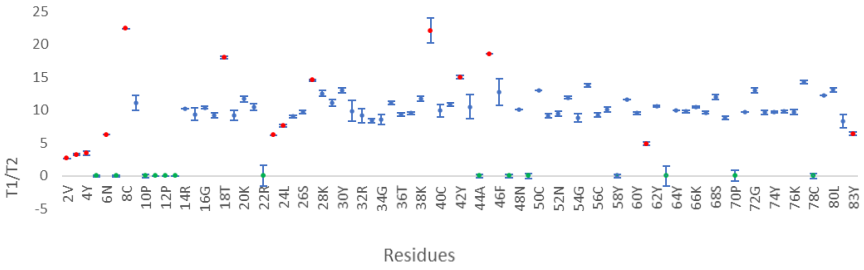


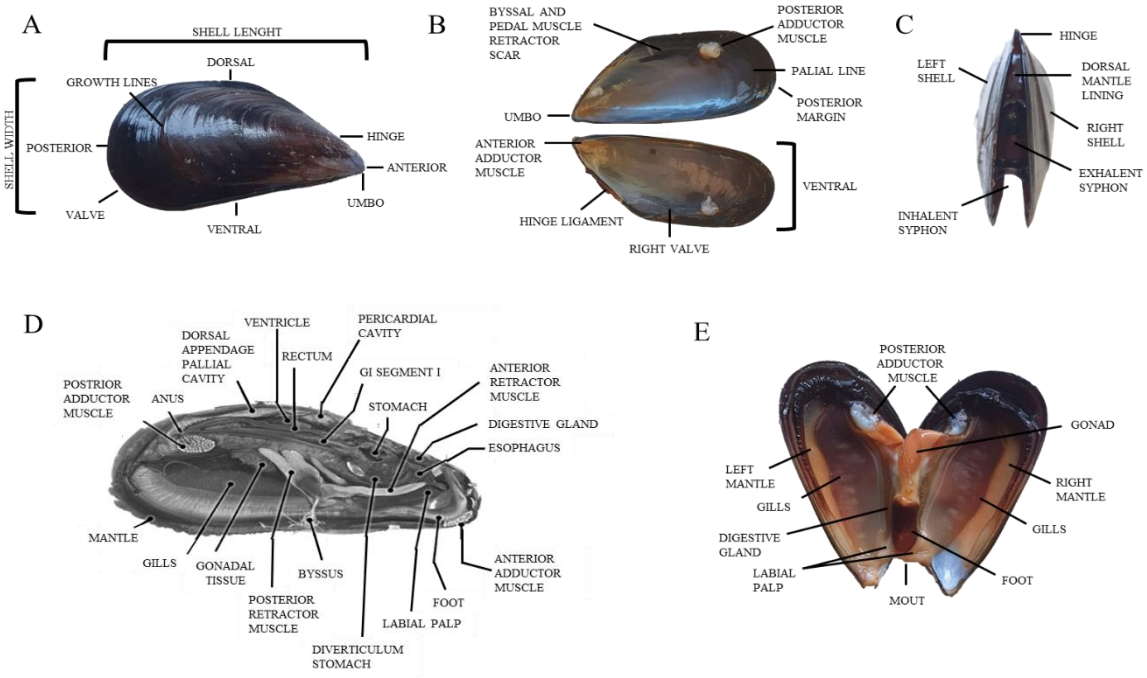
Figure 3.32 T1/T2 ratio measured at 800 MHz. In green residues for which no HeteroNOE was calculated due to the presence of peaks overlap (exception is made for proline no observable in a ^{15}N ^{1}H HSQC spectrum): 5P, 7P, 10P, 11Y, 12P, 13C, 22R, 44A, 47P, 58Y, 63P, 70P and 78C. In red those residues whose T1/T2 ratio value is outside the observed range: 2V, 3Y, 4Y, 6N, 8C, 18T, 23G, 24L, 27Y, 39N, 42Y, 45C, 61G, 83Y.

4. Conclusions

Mussel foot proteins are, among many adhesive proteins, those that have attracted the attention of many reseachers in the last decades, due to its ability to firmly adher in water surfaces remaining tough and durable in time. Despite many efforts have been made to develop biomaterials, capable to mimic exactly the adhesive properties of *mfps*, there are still many questions that remain unresolved and can be able to explain how mussel foot proteins perform its adhesive abilities in wet environment. The knowledge of their three-dimensional structure could provide us with useful informations on how they are able to exert their adhesive properties. In this work, it is presented for the first time the three-dimensional structure of recombinant Pvfp-5β, the first adhesive protein of *Perna viridis* mussel species to be secreted by and initiate interaction with water surface. An *ad hoc* protocol was optimized to produce recombinant Pvfp-5β, which allowed for the first time to produce the protein in its no Dopa form. Pvfp-5β was found to be monomeric and with an ellipse shape, which consists mainly of random coil and two β-sheets, stabilized by five disulfide bridges determined by mass spectrometry. The presence of the disulfide bridges led protein to be more rigid, than what can be expected from protein with high level of random coil structure. Only two of the twelve cysteines are not involved in disulfide bridge due to their distance. Pvfp-5β structure shares many structural similarities with two tandem EGF motifs. A single EGF motif is indeed mainly composed by three disulfide bridges, which hold together two β-strands structure. The 20 NMR structure ensemble of Pvfp-5β generated in this research work, still presents few violations occurring up to 0.5 Å, which cause a non- overlapping of structures in the N-terminal region. However, upper bound violations are still refining. Relaxation studies indicate that overall Pvfp-5β is well ordered, even if fluctuations at ps-ns time scale occur for residues at the N- terminal side. This could explain the lack of good convergence of the structure ensemble in this region. The internal motions occurring in N-terminal region can also explain the reason why the sixth disulfide bridge is not formed and the two cysteines remain decoupled. Few residues in the central region display T1/T2 ratio values outside of the average of entire chain. This could be related with internal motion at µs-ms time scale, but more investigations are still needed to disclosure this point. Although the three-dimensional structure of Pvfp5β is still under refinement, it was possible to build hypotheses of how Pvfp-5β is able to exploit its adhesive behavior with wet surfaces. In

particular, the high presence of basic amino acid residues, positively charged, close to tyrosine residues, seem to influence and be fundamental in the interaction of the protein with surfaces. These residues can be involved in saline bridges, thus they result in attracting surface's charges toward the protein. Tyrosine residues get then close to the surface, and can exploit their hydroxyl group and aromatic ring to performe hydrogen bonds, π-π and π-cation interactions with surface, resulting in enhancing protein adhesion. At basic pH value, plaque solidification occurs in mussel foot, which is justify by covalent interaction with surface, and could be explained also with higher nucleophilic reactivity of lysine and arginine. This results tyrosine residues, close to arginine and lysine residues, are involved also in cohesion processes, due to the proximity to electophilic groups with which initiate a covalent bond, exploiting their hydroxyl groups. Adhesion and cytotoxity tests, performed on Pvfp-5β by collaboration with Dr Rosa Passantino of IBF-CNR, proved that Pvfp-5β retains considerable adhesion properties also in the absence of Dopa. This suggest that Dopa is not crucial and that adhesion is not solely depending on the presence of Dopa. Taking all these data into consideration, Pvfp-5β could be considered as template for a development of a potential bioadhesive. This process can be shorten by deep analysis of the three-dimensional structure of Pvfp-5β, which is now available thank to the work perfomed during this PhD research project.

ANNEX I: Mussel Anatomy



A. Shell of blue mussel. The two simmetric shaped shell valves that protect the mollusc, are joined together by a hinge. On the shell valve we can distinguish different regions: an anterior region, a more rounded posterior region, a dorsal region, and a ventral region. Shell is formed by three layers: i) the periostracum, the external layer, formed by a thin layer of proteic nature, (conchiolin); ii) the ostracum, the second layer, composed of carbonate calcium prism; iii) the ipostracum, or pearly, the third inner layer, which is secreted by the mantle epidermis and is made up of parallel stripes of calcium carbonate. The streaks that we observe, on the external surface of the shell, highlight the animal's growth. **B.** Inner shell of blue mussel. The inside of the shell is composed by: i) an elastic ligament connects the two valves, which are kept closed or opened by the contraction of the anterior adductor muscle; ii) the posterior adductor muscle; iii) a palial cavity. **C.** Oblique dorsal view of blue mussel. Sea water circulates in the palial cavity, it enters through the ventral opening of the mantle, incurrent siphon, and it exits through the excurrent siphon. **D.** Longitudinal micro-CT section in 3D at the level of the heart of a critical point dried blue mussel (Mytilus edulis). Image taken from *Sci Rep*, 2020, **10,** 6773. **E.** Ventral view of blue mussel. The soft body of the

mytilus is wrapped by the mantle. The mantle is divided into two big lobes that surround the visceral mass and it delimits the palial cavity. On the mantle surface, we can see the adductor muscles, anterior and posterior, the retractor muscle, the foot and the byssus. Muscles are microfiltrating organisms and their big gills on both sides of their visceral mass, play a double role for breathing and for nutrition. The gills are characterized by a double row of long parallel filaments that permit to guide the plancton towards the mouth. The filaments, in the inside, are covered by thousands of lashes, whose pulsing permits the circulation of water and breathing exchange. At the base of the gills we can find the kidneys of a dark red color. The muscles' digestive system comprehends the mouth, surrounded by four flat labial palps, the esophagus, the stomach, the hepatopancreas and the intestine. Muscles are also provided of a bilateral symmetry nervous system, formed by three couples of ganglions connected to each other: a couple of cerebral ganglions, a couple of pleural ganglions, and a couple of visceral ganglions, near the posterior adductor muscle. Bivalve mussels have an open circulatory system. The heart is found in the dorsal position, surrounded by the pericardium, and it is made up by two atria and by one ventricle. The ventricle pushes the hemolymph into the anterior and posterior aortas, from here it ends up in irrigating the organs, going through kydneys and gills, before returning to the heart. In the central position, the foot is visible, muscular and of a brownish color, made by a longitudinal posterior ventral groove.

ANNEX II.1: DNA sequence of plasmid phTt-Pvfp5β

Highlighted in grey: Start codon
Highlighted in yellow: His-tag
Highlighted in cyan: Thioredoxin
Highlighted in red: TEV cleavage site
Highlighted in purple: Pvfp-5β
Highlighted in green: Stop codon

CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAG
AGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGC
CGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGA
AAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTG
GCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGC
CCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTG
CCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTG
CACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCA
GGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCT
CTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGC
GTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAG
TTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAA
TTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACC
ATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGAT
GGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATTT
CGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACC
ATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTC
TCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAA
CCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATG
CAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
TAAGTTAGCTCACTCATTAGGCACAATTCTCATGTTTGACAGCTTATCATCGACTGCAC
GGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGG
TCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTT
TTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAA
TCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAG
CCAGTCCGTTTAGGTGTTTTCACGAGCAATTGACCAACAAGGACCATAGATT**ATGAAAT**
CTTCTCACCATCACCATCACCAT**GGTTCTTCTATGAGCG**ATAAAATTATTCACCTGACT
GACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCCTCGTCGATTTCTG
GGCAGAGTGGTGCGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACG
AATATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCG
CCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGC
GGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACC
TGGCCGGGATCGAG**GAAAACCTGTACTTTCAAGGC**GTGTACTATCCGAACCCGTGTTCT
CCATACCCGTGCCGCAACGGTGGTACCTGTAAGAAACGTGGTCTGTACTCTTATAAATG
TTATTGCCGTAAAGGTTATACCGGCAAGAACTGTCAATACAACGCATGTTTTCCGAACC
CGTGTCTGAACGGTGGCACTTGCGGTTACGTTTACGGTTATCCGTATTACAAGTGTAGC
TGCCCATACGGTTATTACGGCAAACAATGCCAACTGAAAAAGTA**TTA**AGAATTCCCTGC
AGGTAATTAAATAAGCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCG
TTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGG

ATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACT
GCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACA
AACTCTTTCGGTCCGTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAA
CATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCA
CCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTT
ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGT
TTCCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGA
CGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGT
ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGT
GCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGG
ACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATC
GTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCT
GTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTC
CCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCT
CGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCT
CGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTA
CACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG
CCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATT
GATTTCCTTAGGACTGAGCGTCAACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATC
CTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTG
GTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAG
AGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGA
ACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCC
AGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGC
GCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCT
ACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGG
AGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGA
GCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC
TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGC
AACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCC
TGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCG
CTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC
CTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATAAGGTGC
ACTGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA
CGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTG
CATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCT
CATCAGCGTGGTCGTGCAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCG
TTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGC
GGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGG
TAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACAT
GCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCA
GAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCAC
AGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGAC
TTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCA
GGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCAT
TCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACG
ATCATGCGCACCCGTGGCCAGGACCCAACGCTGCCCGAAATT

ANNEX II.2: NMR Experiments pulse sequences

All images in this Annex II are taken from Index of /guide/tutorials/sequences (csic.es).

$^{1}H^{15}N$ HSQC

The two-dimensional $^{1}H^{15}N$ HSQC pulse sequence (Fig. II.1) starts with a pre-scan period, usually defined by a long recycle delay, d1, to allow the recovery of the ^{1}H magnetization to a pre-equilibrium state just before to start the sequence. After the initial ^{1}H excitation, by applying a ^{1}H 90 degree pulse, a fixed period, d26, optimized to 1/4J(NH) is applied to allow evolution of heteronuclear J(NH) coupling costant. Simoultaneous ^{1}H and ^{15}N 180 degree pulse is applied to refocus chemical shift evolution. At the end of the fixed period, d26, antiphase ^{1}H magnetization, with respect to ^{15}N, is obtained. A double 90 degree pulse both to ^{1}H and ^{15}N is applied to transfer magnetization from ^{1}H to ^{15}N. The above sequence described is known as Insensitive Nuclei Enhancement by Polarization Transfer (INEPT). A water-selective 90º pulse creates transverse magnetization for the water whereas other signals are unaffected. As a general approach, the watergate block is usually applied during the retro-INEPT block of the standard HSQC pulse sequence in order to improve solvent suppression in aqueous samples. The selective 90 degree pulse applied on the water resonance is defined by the pulse lenght for 90 degree shaped pulse (p11) and the power level for 90 degree shaped pulse (sp1). The gradients G1, G2 and G3 dephases water signal. During the evolution time, t1, the ^{15}N chemical shift and ^{1}H-^{15}N coupling evolve. The ^{1}H 180° pulse in the center of the evolution time refocuses the coupling and as a result decouples protons in the F1 (^{15}N) domain of the spectrum. Watergate is also combined with the water flip-back pulse which ensures that the water magnetization is little perturbed and oriented along the +z axis during most of the experiment, especially just prior to acquisition, to minimize the saturation of water. The water flip-back pulse is specified by the pulse lenght for 90 degree shaped pulse (p29) and the power level for 90 degree shaped pulse (sp11). During the retro-INEPT sequence, optimized to 1/4J(NH), ^{15}N magnetizzation is transferred back to ^{1}H nuclei by applying a 90 degree pulse on ^{15}N. Simoultaneous 180 degree pulse, applied to ^{1}H and ^{15}N, refocus the ^{1}H chemical shift evolution. The pulse sequence-end

was achieved by Globally optimized Alternating-phase Rectangular Pulses (GARP) to suppress ^{15}N-^{1}H heteronuclear scalar coupling during acquisition.

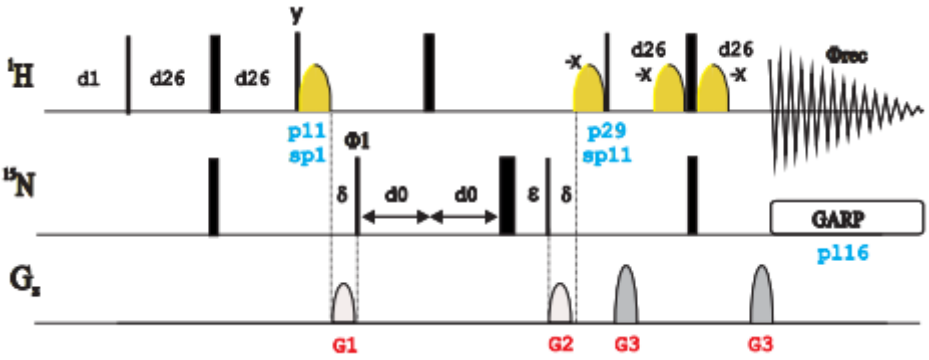


Figure II.1 Pulse sequence of phase sensitive 2D ^{1}H^{15}N HSQC experiment, using water flip-back and watergate suppression (hsqcfpf3gpphwg). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses.

HNCACB

The pulse sequence of the three-dimensional HNCACB experiment (Fig. II.2) consist of a relaxation period, d1, on ^{1}H channel followed by a standard INEPT sequence, optimized to 1/4J(NH), to transfer magnetization from ^{1}H to ^{15}N via J(NH). The watergate suppression pulse is applied by a selective 90 degree pulse on the water resonance and is defined by the pulse lenght for 90 degree shaped pulse (p11) and the power level for 90 degree shaped pulse (sp1) resulting in restore the water magnetization on z axis. The gradient G1 dephases water signal. A second INEPT sequence on ^{15}N, optimized to 1/4J(NCa), follows to transfer magnetization from ^{15}N to ^{13}CA. The ^{1}H 90 degree pulse applied on y, at the beginning of the DIPSI-2 sequence, allow to focus the magnetization water on x axis and lock it during 1H decoupling. A 90 degree pulse on ^{13}Cαβ transfer magnetization from ^{13}CA to ^{13}CB followed by the ^{13}Cαβ chemical shift evolution during the evolution time t1. Decoupling In the Presence of Scalar Interactions (DIPSI-2) pulse sequence is specifically employed to decouple protons. Scalar coupling of the ^{13}C carbonyl carbons and ^{15}N are removed by a selective 180° pulse during the ^{13}Cαβ evolution. A 90 degree pulse on ^{13}Cαβ transfer back the magnetization from ^{13}CB to ^{13}CA carbons. Water magnetization is restored on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. All the transverse magnetization are dephased by the gradient G2. Magnetization is transferred back from ^{13}CA to ^{15}N. A ^{1}H 90 degree pulse on y axis, applied at the beginning of the second DIPSI-2 sequence, rotate the magnetization water

on x axis during ^{1}H decoupling. The ^{15}N chemical shift evolution occur during the constant period, t2, during which DIPSI-2 pulse sequence is applied on protons. Water magnetization is restored again on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. The decoupling of protons is turn off for a time d21, allowing the ^{15}N spins to become antiphase. All the transverse magnetization are dephased by the gradient G3 and G4. The magnetization is transferred back to the NH protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{15}N by GARP pulse.

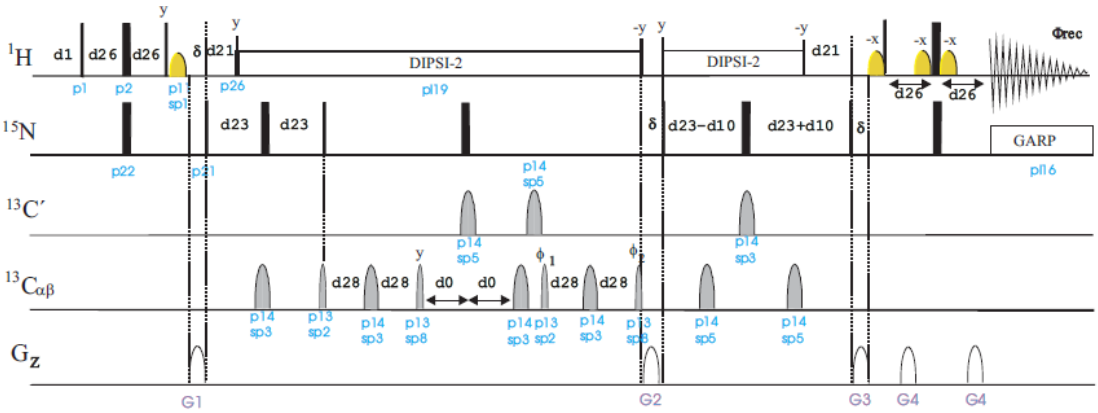


Figure II.2. Pulse sequence of 3D HNCACB experiment, using watergate suppression (hncacbgpwg3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses.

HN(CO)CACB

The pulse sequence of three-dimensional HN(CO)CACB experiment (Fig. II.3) start with a relaxation period, d1, to achieve a pre-equilibrium state before the INEPT scheme, optimized to 1/4J(NH), is applied on ^{1}H channel, during which the magnetization is transferred from ^{1}H to ^{15}N via J(NH). A watergate pulse follows by applying a 90 degree shaped pulse (p11) and a power level for 90 degree shaped pulse (sp1) on 1H channel. The gradient G1 dephases water signal. The ^{1}H 90 degree pulse applied on y, at the beginning of the DIPSI-2 sequence, allow to focus the magnetization water on x axis and lock it during ^{1}H decoupling. A second INEPT sequence on ^{15}N, optimized to 1/4J(NCO), follows to transfer magnetization from ^{15}N to ^{13}CO. Follows fixed evolution delay, d22, to achieve antiphase ^{13}CO magnetization with respect to $^{13}C\alpha$ via $^{1}J(CoCa)$. The magnetizzation is transferred from ^{13}C carbonyl carbons to ^{13}CA by applying a 90 degree

pulse on ^{13}C carbonyl carbons. A 90 degree pulse on $^{13}C\alpha\beta$ transfer magnetization from ^{13}CA to ^{13}CB. During the $^{13}C\alpha\beta$ chemical shift evolution time, t1, scalar coupling of the ^{13}C carbonyl carbons, ^{15}N and protons are removed by a selective 180° pulse and DIPSI-2 scheme respectivelly. A 180 degree pulse on $^{13}C\alpha\beta$ transfer back the magnetization from ^{13}CB to ^{13}CA carbons. Fixed evolution delay is applied to achieve antiphase ^{13}CA magnetization with respect to ^{13}CO via 1J(CA,CO) followed by magnetization transfer to ^{13}CO by simultaneous 90 degree pulses on $^{13}C\alpha\beta$ and ^{13}CO. Another fixed delay occur to allow antiphase ^{13}C carbonyl carbons magnetization with respect to ^{15}N via 1J(N,CO) followed by magnetization transfer to ^{15}N by simultaneous 90 degree pulses on ^{15}N and ^{13}CO. Water magnetization is restored on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. All the transverse magnetization are dephased by the gradient G2. A ^{1}H 90 degree pulse on y axis, applied at the beginning of the second DIPSI-2 sequence, rotate the magnetization water on x axis during 1H decoupling. The ^{15}N chemical shift evolution occur during the constant period t2, during which DIPSI-2 pulse sequence, ^{13}CO and $^{13}C\alpha\beta$ decoupling are applied. Water magnetization is restored again on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. The decoupling of protons is turn off for a time d21, allowing the ^{15}N spins to become antiphase. All the transverse magnetization are dephased by the gradient G3 and G4. The magnetization is transferred back to the NH protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{15}N by GARP pulse.

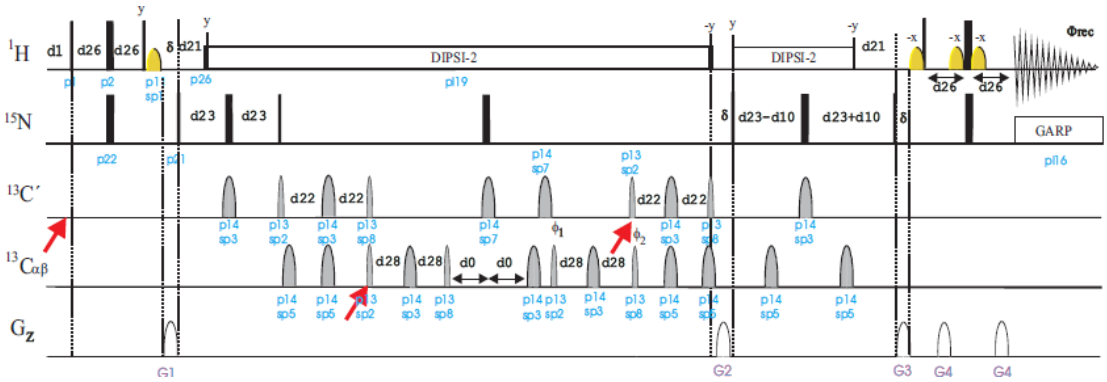


Figure II.3. Pulse sequence of 3D HN(CO)CACB experiment, using watergate suppression (hncocacbgpwg3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

HN(CO)CA

Pulse sequence for three-dimensional HN(CO)CA experiment (Fig. II.4) uses at first a relaxation period d1, followed by an INEPT sequence to transfer magnetization from amide protons to ^{15}N. A watergate pulse follows by applying a 90 degree shaped pulse (p11) and a power level for 90 degree shaped pulse (sp1) on 1H channel. A second INEPT sequence on ^{15}N, optimized to 1/4J(NCO), follows to transfer magnetization from ^{15}N to ^{13}CO. The 1H 90 degree pulse applied on y, at the beginning of the DIPSI-2 sequence, allow to focus the magnetization water on x axis and lock it during 1H decoupling. Follows fixed evolution delay, d22, to achieve antiphase ^{13}CO magnetization with respect to $^{13}C\alpha$ via $^{1}J(CoCa)$. The magnetizzation is transferred from ^{13}C carbonyl carbons to ^{13}CA by applying a 90 degree pulse on ^{13}C carbonyl carbons. The $^{13}C\alpha$ carbons and protons are decoupled by 180 degree pulse and DIPSI-2 scheme respectivelly during the ^{13}C carbonyl carbons chemical shift evolution. During the $^{13}C\alpha$ chemical shift evolution time, t1, scalar coupling of the ^{13}C carbonyl carbons, ^{15}N and protons are removed by a selective 180° pulse and DIPSI-2 scheme respectivelly. The magnetization is transferred back to ^{13}CO by simultaneous 90 degree pulses on $^{13}C\alpha$ and ^{13}CO. A fixed delay occur on ^{13}C channel to allow antiphase ^{13}C carbonyl carbons magnetization with respect to ^{15}N via $^{1}J(N,CO)$ followed by magnetization transfer to ^{15}N by simultaneous 90 degree pulses on ^{15}N and ^{13}CO. Water magnetization is restored on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. All the transverse magnetization are dephased by the gradient G2. A ^{1}H 90 degree pulse on y axis, applied at the beginning of the second DIPSI-2 sequence, rotate the magnetization water on x axis during ^{1}H decoupling. The ^{15}N chemical shift evolution occur during the constant period t2, during which DIPSI-2 pulse sequence, ^{13}CO and $^{13}C\alpha$ decoupling are applied. Water magnetization is restored again on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. The decoupling of protons is turn off for a time d21, allowing the 15N spins to become antiphase. All the transverse magnetization are dephased by the gradient G3 and G4. The magnetization is transferred back to the NH protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{15}N by GARP pulse.

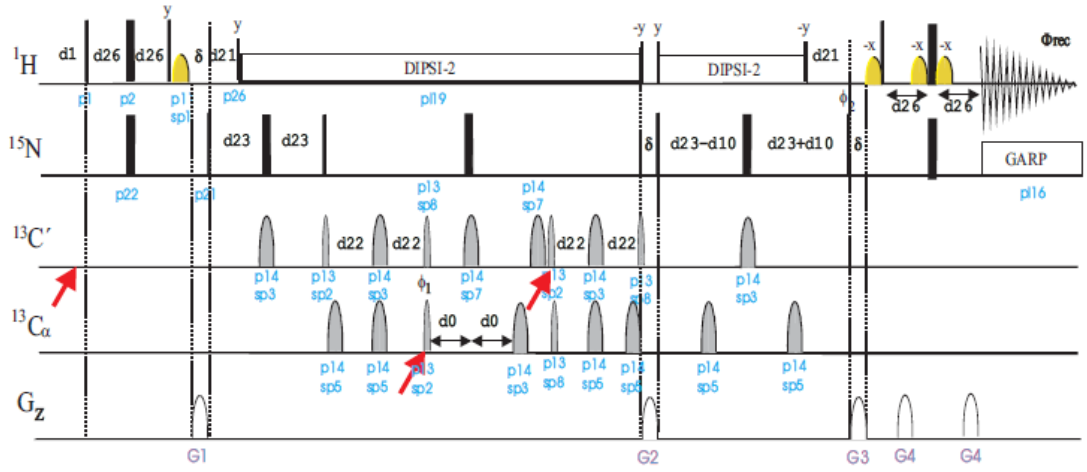


Figure II.4. Pulse sequence of 3D HN(CO)CA experiment, using watergate suppression (hncocagpwg3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

HNCA

In three-dimensional HNCA experiment the magnetization rises from amide protons (Fig. II.5) is transferred, after a relaxation delay, on ^{15}N nuclei by an INEPT sequence, optimized to 1/4J(NH). A watergate pulse follows by applying a 90 degree shaped pulse (p11) and a power level for 90 degree shaped pulse (sp1) on ^{1}H channel. The application of pulse p21 on ^{15}N establishes antiphase nitrogen magnetization which is refocus during the delay d21. The ^{1}H 90 degree pulse applied on y, at the beginning of the DIPSI-2 sequence, allow to focus the magnetization water on x axis and lock it during ^{1}H decoupling. The gradient G1 dephases water signal. A second INEPT sequence on ^{15}N, optimized to 1/4J(NCa), follows to transfer magnetization from ^{15}N to ^{13}CA. During the ^{13}Cα chemical shift evolution time, t1, scalar coupling of the ^{13}C carbonyl carbons, ^{15}N and protons are removed by a selective 180° pulses and DIPSI-2 scheme respectively. Water magnetization is restored on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. All the transverse magnetization are dephased by the gradient G2. A ^{1}H 90 degree pulse on y axis, applied at the beginning of the second DIPSI-2 sequence, rotate the magnetization water on x axis during ^{1}H decoupling. Follows the magnetization transfer from ^{13}Cα to ^{15}N with nitrogen chemical shift evolution taking place during the constant time delay t2, during which DIPSI-2 pulse sequence, ^{13}CO and ^{13}Cα decoupling are applied. The decoupling of protons is turn off for a time d21, allowing the ^{15}N spins

to become antiphase. All the transverse magnetization are dephased by the gradient G3 and G4.The magnetization is transferred back to the NH protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{15}N by GARP pulse.

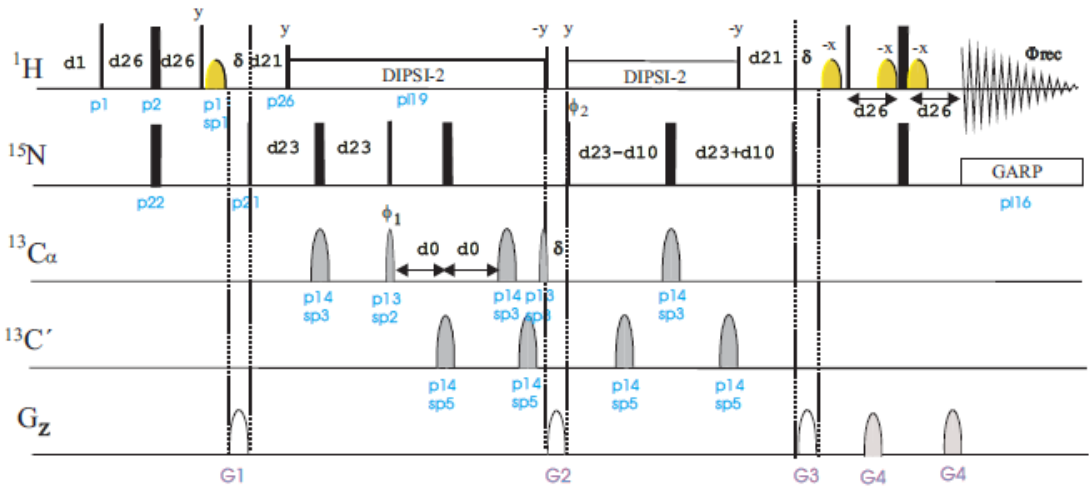


Figure II.5. Pulse sequence of 3D HNCA experiment, using watergate suppression (hncagpwg3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

HNCO

In three-dimensional HNCO experiment the magnetization rises from amide protons (Fig. II.6) is transferred, after a relaxation delay, on ^{15}N nuclei by an INEPT sequence optimized to 1/4J(NH). A watergate pulse follows by applying a 90 degree shaped pulse (p11) and a power level for 90 degree shaped pulse (sp1) on ^{1}H channel. The gradient G1 dephases water signal A second INEPT sequence on ^{15}N, optimized to 1/4J(NCO), follows to transfer magnetization from ^{15}N to ^{13}CO. The ^{1}H 90 degree pulse applied on y, at the beginning of the DIPSI-2 sequence, allow to focus the magnetization water on x axis and lock it during ^{1}H decoupling. During the ^{13}CO chemical shift evolution time, t1, scalar coupling of the $^{13}C\alpha$, ^{15}N and protons are removed by a selective 180° pulses and DIPSI-2 scheme respectively. Water magnetization is restored on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. All the transverse magnetization are dephased by the gradient G2. A ^{1}H 90 degree pulse on y axis, applied at the beginning of the second DIPSI-2 sequence, rotate the magnetization water on x axis during ^{1}H decoupling. Follows the magnetization transfer from ^{13}C carbonyl carbons to ^{15}N with

nitrogen chemical shift evolution taking place during the constant time delay t2, during which DIPSI-2 pulse sequence, $^{13}C\alpha$ and ^{13}CO decoupling are applied. Water magnetization is restored again on z axis applying a ^{1}H 90 degree pulse on -y at the end of DISPI-2 pulse. The decoupling of protons is turn off for a time d21, allowing the ^{15}N spins to become antiphase. All the transverse magnetization are dephased by the gradient G3 and G4. The magnetization is transferred back to the NH protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{15}N by GARP pulse.

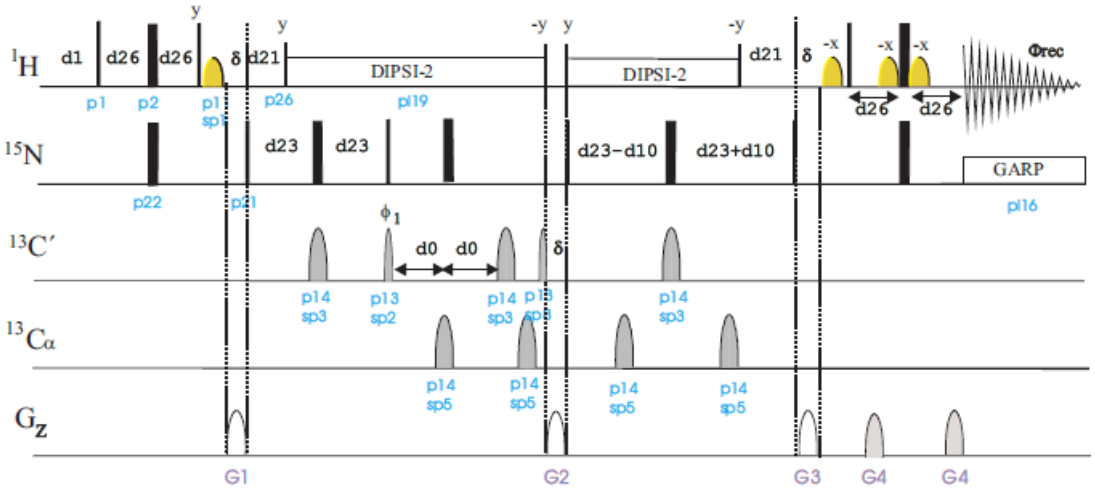


Figure II.6. Pulse sequence of 3D HNCO experiment, using watergate suppression (hncogpwg3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

HBHA(CO)NH

The pulse sequence of three-dimensional HBHA(CO)NH experiment (Fig. II.7) evolves with the transverse magnetization of Hα and Hβ during the evolution period, t1, due to the initial ^{1}H 90 degree pulse. At the same time the magnetization becomes antiphase with respect to ^{13}C. It is refocused during delay d3. Follows the magnetization transfer to the close $^{13}C\alpha\beta$ by applying a ^{1}H 90 degree pulse. Fixed evolution delay, d22, is applied to achieve antiphase $^{13}C\alpha$ magnetization with respect to $^{13}C\beta$ via 1J(CA,CB), during which protons decoupling is applied by DIPSI-2 sequence, followed by magnetization transfer from $^{13}C\beta$ to $^{13}C\alpha$ by applying a 90 degree pulse. Another fixed evolution delay, d23, is applied to achieve antiphase $^{13}C\alpha$ magnetization with respect to ^{13}CO via 1J(CACO) followed by magnetization transfer to ^{13}CO by applying 90 degree pulses on $^{13}C\alpha\beta$ and

^{13}CO. To achieve ^{13}CO antiphase, with respect to ^{15}N, a fixed delay d27 is applied thanks to the 1J(NCO), followed by transfer magnetization to ^{15}N by applying 90 degree pulses on ^{13}CO and ^{15}N. The gradients G1, G2 and G3 applied dephases water signal. During the costant time evolution of ^{15}N magnetization, t2, selective decoupling is applied on the ^{13}Cαβ region and on protons by DISPSI-2 sequence. The magnetization of the amide proton is detected during the time t3 while the ^{15}N is decoupled by GARP pulse.

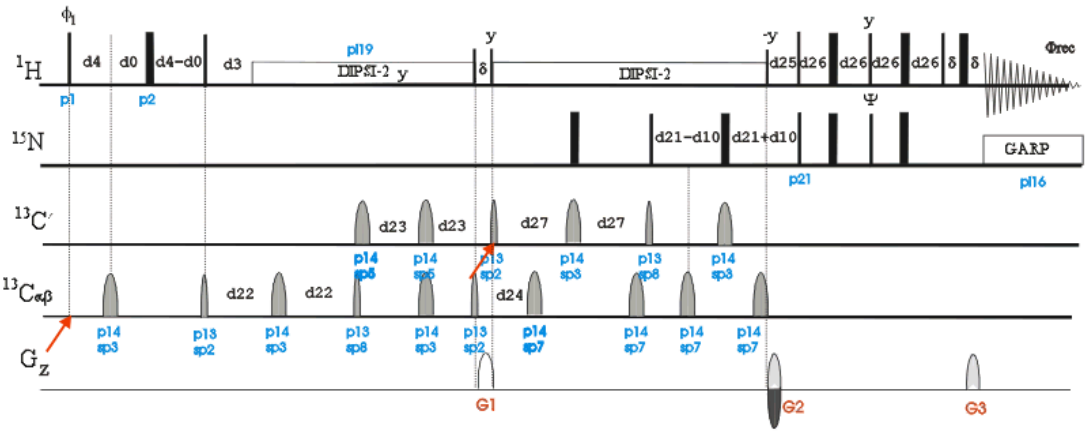


Figure II.7. Pulse sequence of 3D HBHA(CO)NH experiment, (hbhaconhgp3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses.

HBHANH

The pulse sequence of three-dimensional HBHANH experiment (Fig. II.8) evolves in the same way as the HBHA(CO)NH. Transverse magnetization of Hα and Hβ occur during the evolution period, t1, due to the initial ^{1}H 90 degree pulse and simoultaneously the magnetization becomes antiphase with respect 13Cαβ. It is refocused during delay d3. Follows the magnetization transfer to ^{13}Cαβ nuclei by applying a ^{1}H 90 degree pulse. Fixed evolution delay, d22, is applied to achieve antiphase ^{13}Cα magnetization with respect to ^{13}Cβ via 1J(CA,CB), during which protons decoupling is applied by DIPSI-2 sequence, followed by magnetization transfer from ^{13}Cβ to ^{13}Cα by applying a 90 degree pulse. Another fixed evolution delay, d27, is applied to achieve antiphase ^{13}Cα magnetization with respect to ^{15}N via ^{1}J(NCA) followed by magnetization transfer to ^{15}N by applying 90 degree pulses on ^{13}Cαβ and ^{15}N.The gradient G1 dephases water signal. Nitrogen chemical shift evolution taking place during the constant time delay t2. During

the costant time evolution of ^{15}N magnetization, selective decoupling is applied on the $^{13}C\alpha\beta$ region, on ^{13}C carbonyl carbons and on protons by DISPSI-2 sequence. The magnetization of the amide proton is detected during the time t3 while the ^{15}N is decoupled by GARP pulse.

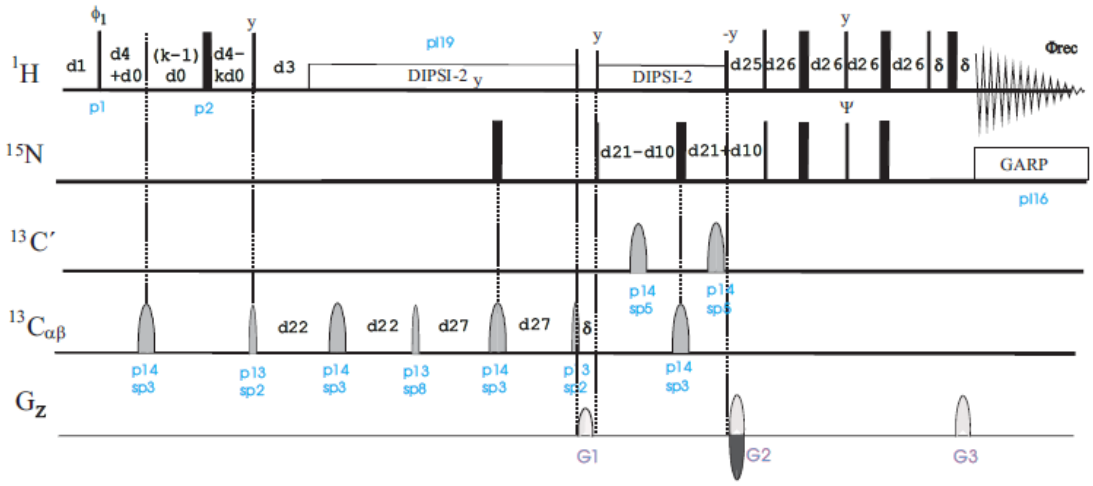


Figure II.8. Pulse sequence of 3D HBHANH experiment, (hbhanhgp3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

$^{1}H^{13}C$ HSQC

The pulse sequence of bidimensional $^{1}H^{13}C$ HSQC experiment (Fig. II.9) starts with an initial ^{1}H excitation, by applying a ^{1}H 90 degree pulse, followed by a fixed period, d4, optimized to 1/4J(CH) to allow evolution of heteronuclear J(CH) coupling costant. Then a trim pulse, p28, is applied to dephase excess magnetization. A double 90 degree pulse both to ^{1}H and ^{13}C is applied to transfer magnetization from ^{1}H to ^{13}C.

During the costant time evolution of 13C magnetization, protons are decoupled by applying a 180 degree pulse on 1H channel. A heteronuclear gradient, Echo/Antiecho-TPP, is applied for the coherence selection. The magnetization is transferred back to ^{1}H nuclei for detection by a modification of the reverse INEPT block to afford improved sensitivity using the PEP (Preservation of Equivalent Pathways) methodology. Proton acquisition is recorded decoupling ^{13}C by GARP pulse.

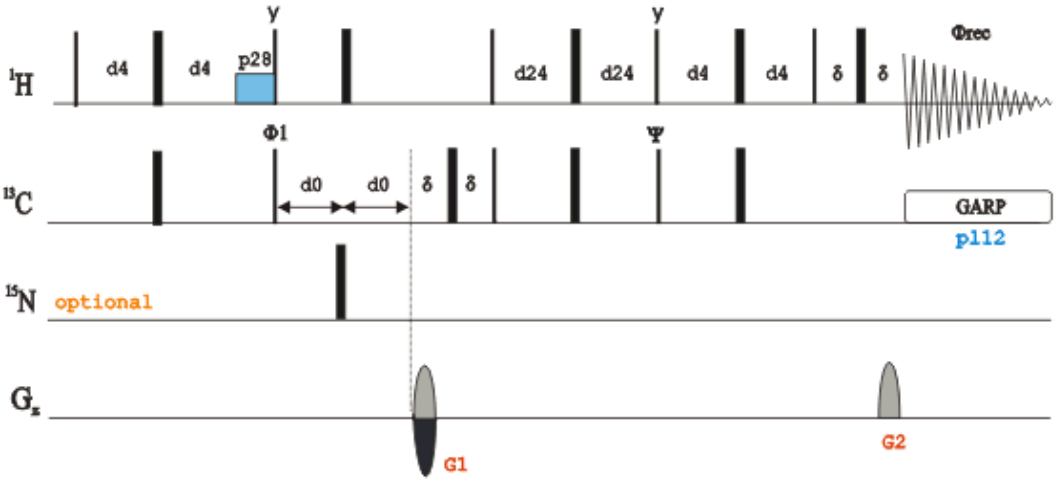


Figure II.9. Pulse sequence of 2D ^{1}H ^{13}C HSQC experiment, using a phase sensitive echo/antiecho-TPPI gradient with sensitivity improvement (hsqcetgpsi). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

HCCH-TOCSY

The pulse sequence of three-dimensional HCCH-TOCSY (Fig. II.10) starts with the magnetization of proton after the ^{1}H 90 degree pulse. Follows ^{1}H chemical shift evolution during the evolution time t1 whereas ^{13}C nuclei are decoupled by applying 180 degree pulse. Fixed evolution delay is applied to achieve antiphase ^{1}H magnetization with respect to ^{13}C through $^{1}J(C,H)$. G1 and G3 gradient ensure that only ^{1}H transverse magnetization is present and water signal is suppressed. The others gradients behave at the same way as the preceding. The magnetization is transferred to ^{13}C nuclei by applying 90 degree pulses on both ^{1}H and ^{13}C channel. ^{13}C chemical shift evolution occur during the evolution time t2 and applying protons decoupling by 180 degree pulse. Fixed evolution delay is applied to achieve antiphase ^{13}C magnetization with respect to ^{1}H via $^{1}J(C,H)$. A 90 degree pulse, on y axis, is applied at the beginning of the DIPSI-3 sequence, rotating ^{13}C magnetization on x axis and lock it during the isotropic mixing sequence (DIPSI-3). DIPSI-3 block ensures transfer magnetization along the ^{13}C side chain via 1J (C,C) belonging to the same spin network. The trim pulse, p28, at the end of isotropic mixing restore ^{13}C magnetization to z axis. A reverse INEPT sequence transfer ^{13}C magnetization back to protons for detection during which a GARP pulse is applied on ^{13}C channel.

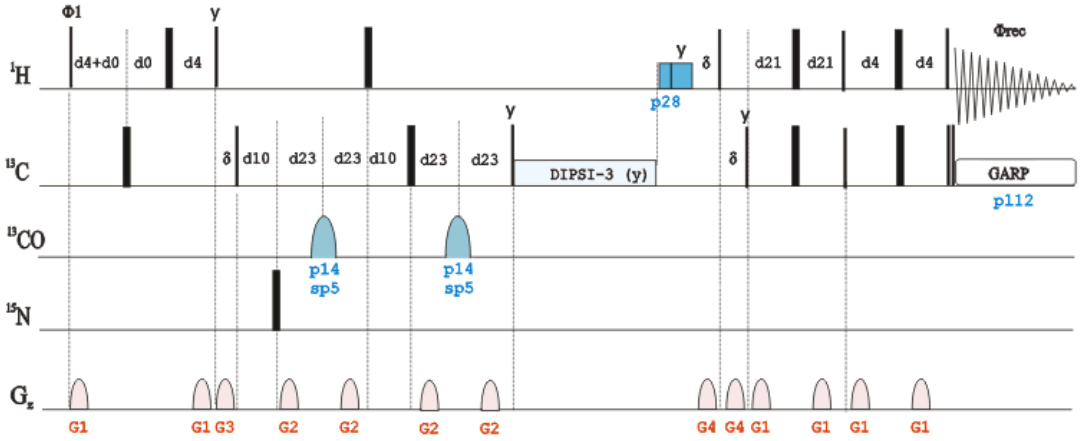


Figure II.10. Pulse sequence of 3D 3D HCCH-TOCSY experiment, (hcchdigp3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

HCCH-COSY

A complementary experiment to HCCH-TOCSY is the HCCH-COSY (Correlated Spectroscopy) experiment which is based on transfer over one ^{13}C-^{13}C bond and correlates neighbouring ^{1}H frequencies only. The pulse sequence of three-dimensional HCCH-COSY (Fig. II.11) it is very similar to HCCH-TOCSY one. It starts with the magnetization of proton after the ^{1}H 90 degree pulse. Follows ^{1}H chemical shift evolution during the evolution time t1 whereas ^{13}C nuclei are decoupled by applying 180 degree pulse. Fixed evolution delay is applied to achieve antiphase ^{1}H magnetization with respect to ^{13}C through $^{1}J(C,H)$. G1 and G3 gradient ensure that only ^{1}H transverse magnetization is present and water signal is suppressed. The others gradients behave at the same way as the preceding. The magnetization is transferred to ^{13}C nuclei by applying 90 degree pulses on both ^{1}H and ^{13}C channel. ^{13}C chemical shift evolution occur during the evolution time t2 and applying protons decoupling by 180 degree pulse. Fixed evolution delay is applied to achieve antiphase ^{13}C magnetization with respect to ^{1}H via $^{1}J(C,H)$. Follows transfer magnetization along the ^{13}C side chain via ^{1}J (C,C) belonging to the same spin network. The trim pulse, p28, at the end of isotropic mixing restore ^{13}C magnetization to z axis. A reverse INEPT sequence transfer ^{13}C magnetization back to protons for detection during which a GARP pulse is applied on ^{13}C channel.

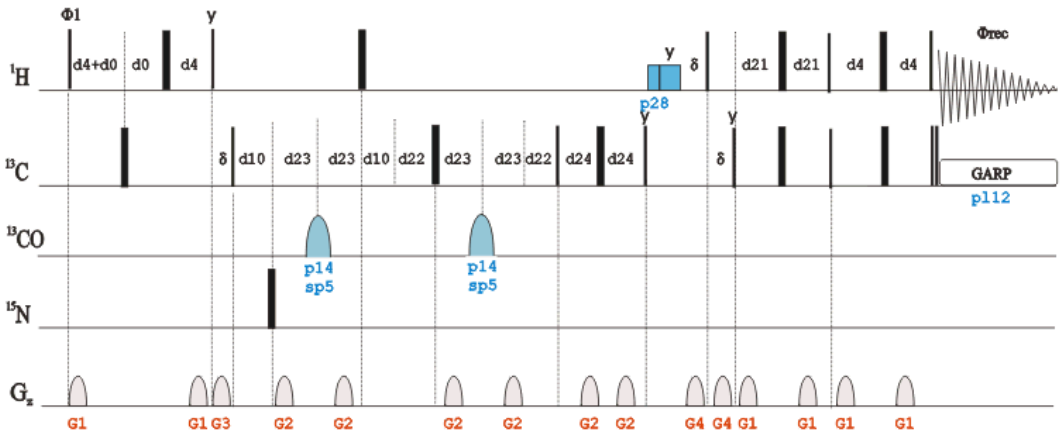


Figure II.11. Pulse sequence of 3D 3D HCCH-COSY experiment, (hcchcogp3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

(HB)CB(CGCD)HD and (HB)CB(CGCD)CEHE

The pulse sequence of two-dimensional (HB)CB(CGCD)HD experiment (Fig. II.12) starts with a trim pulse and a relaxation period, d1, to restore the pre-equilibrium state. The first gradient applied, G1, after a ^{13}C 90 degree pulse remove carbon magnetization to make sure that signals are originated only from protons. Gradients are used in the sequence to suppress artifacts and to remove water signal. Then magnetization is transferred, through an INEPT sequence, to ^{13}Cβ nuclei. Transverse magnetization of ^{13}Cβ occur during the evolution period t1 and protons decoupling. Follows the magnetization transfer from ^{13}Cβ to ^{13}Cγ by applying a 90 degree pulse on ^{13}Cβ. Then the magnetization is transferred from ^{13}Cγ to ^{13}Cδ nuclei. The magnetization is transferred back to protons for detection during $^{13}C_{aro}$ GARP pulse. The selective 90 degree pulse applied on the water resonance is defined by the pulse lenght for 90 degree shaped pulse (p11) and the power level for 90 degree shaped pulse (sp1).

The two-dimensional (HB)CB(CGCD)CEHE pulse sequence (Fig. II.13) is similar to (HB)CB(CGCD)HD experiment, starting with a trim pulse and a relaxation period, d1, to restore the pre-equilibrium state. The first gradient applied, G1, after a ^{13}C 90 degree pulse remove carbon magnetization to make sure that signals are originated only from protons. Gradients are used in the sequence to suppress artifacts and to remove water signal. Then magnetization is transferred, through an INEPT sequence, to ^{13}Cβ nuclei.

Transverse magnetization of $^{13}C\beta$ occur during the evolution period t1 and protons decoupling. Follows the magnetization transfer from $^{13}C\beta$ to $^{13}C\gamma$ by applying a 90 degree pulse on $^{13}C\beta$. Then the magnetization is transferred from $^{13}C\gamma$ to $^{13}C\delta$ and from $^{13}C\delta$ to $^{13}C\varepsilon$ by applying a 90 degree pulse. The magnetization is transferred back to protons for detection during $^{13}C_{aro}$ GARP pulse. The selective 90 degree pulse applied on the water resonance is defined by the pulse lenght for 90 degree shaped pulse (p11) and the power level for 90 degree shaped pulse (sp1).

Figure II.12. Pulse sequence of 2D (HB)CB(CGCD)HD experiment, (hbcbcgcdhdgp). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

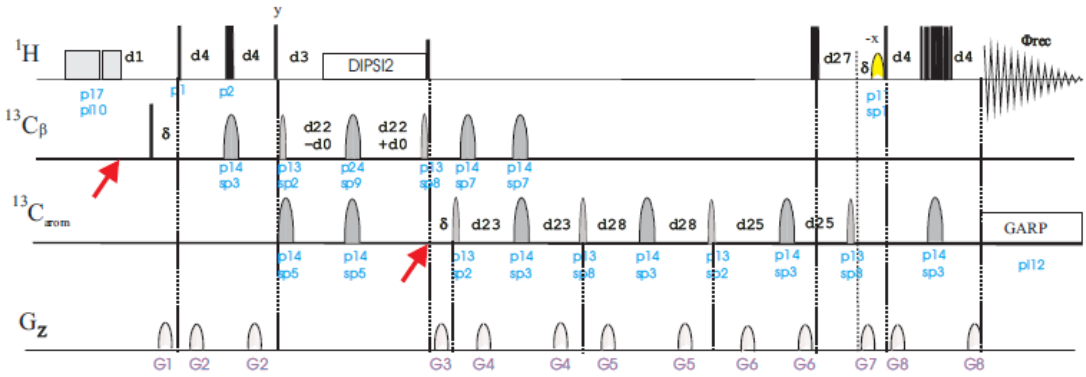


Figure II.13. Pulse sequence of 2D (HB)CB(CGCD)CEHE experiment, (hbcbcgcdcehegp). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

^{15}N NOESY-HSQC

The pulse sequence of the 3D ^{15}N NOESY-HSQC experiment (Fig. II.14) starts with protons magnetization by applying a ^{1}H 90 degree pulse. The chemical shifts of protons

are labeled in t1 evolution during which ^{15}N nuclei are decoupled and there is a noesy-transfer from protons to protons thanks to the mixing time period d8. Magnetization transfer occurs from the proton attached to ^{15}N through INEPT sequence, whereas a trim pulse is applied on ^{1}H channel. The ^{15}N chemical shifts are labeled in t2 evolution followed by a heteronuclear gradient, Echo/Antiecho-TPP for the coherence selection. The magnetization is transferred back to the NH protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{15}N by GARP pulse.

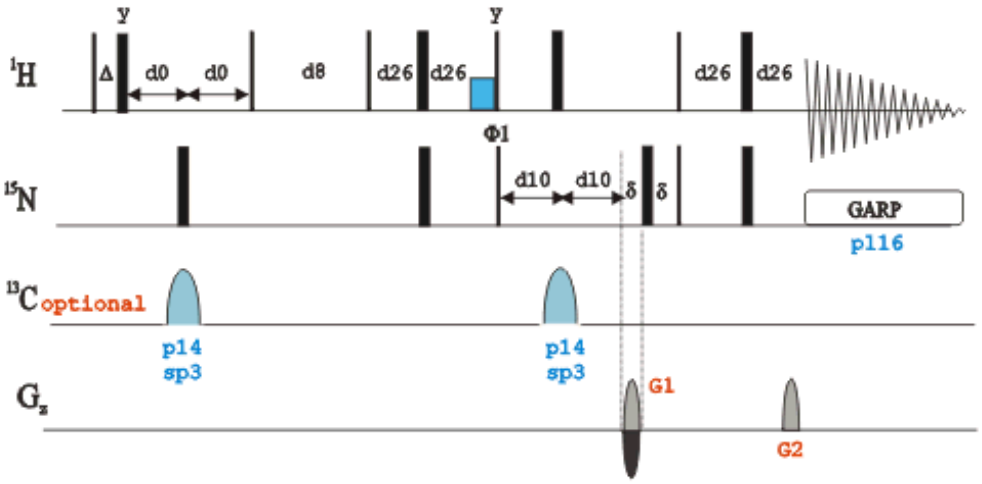


Figure II.14. Pulse sequence of 3D ^{15}N NOESY-HSQC experiment, using a phase sensitive echo/antiecho-TPPI gradient (noesyhsqcetf3gp3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

^{13}C NOESY-HSQC

Pulse sequence for three-dimensional 3D ^{13}C NOESY-HSQC (Fig. II.15) uses at first a relaxation period, d1, to restore the pre-equilibrium state. Follows protons magnetization by applying a ^{1}H 90 degree pulse. The chemical shifts of protons are labeled in t1 evolution during which ^{13}C nuclei are decoupled by applying a 180 degree shaped pulse (p14) and the power level for 180 degree shaped pulse (sp3). A noesy-transfer occurs from protons to protons thanks to the mixing time period d8. Magnetization transfer occurs from the proton attached to ^{13}C through INEPT sequence, whereas a trim pulse is applied on ^{1}H channel. The ^{13}C chemical shifts are labeled in t2 evolution followed by a heteronuclear gradient, Echo/Antiecho-TPP for the coherence selection.

The magnetization is transferred back to the amide protons for detection in t3 by a retro-INEPT block and proton acquisition is recorded decoupling ^{13}C nuclei by GARP pulse.

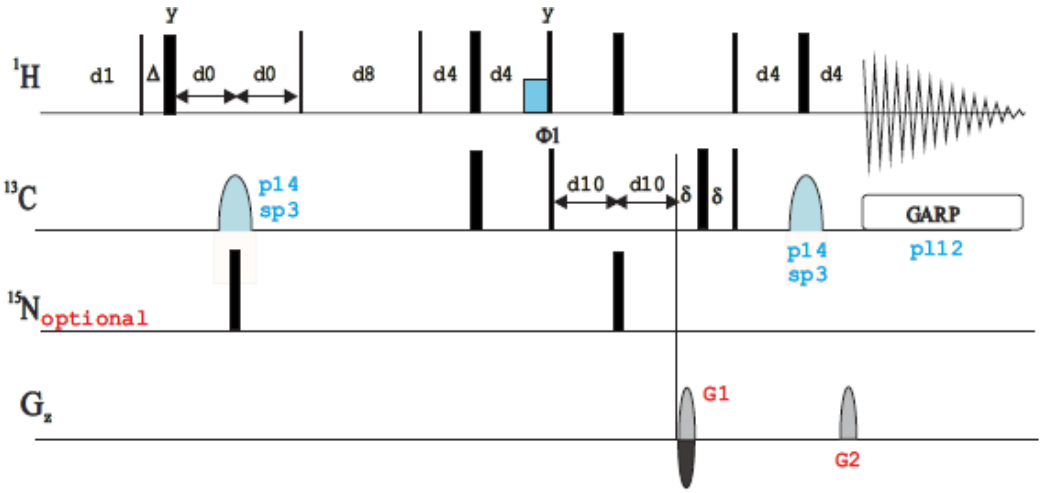


Figure II.15. Pulse sequence of 3D ^{13}C NOESY-HSQC experiment, using a phase sensitive echo/antiecho-TPPI gradient (noesyhsqcetgp3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

T1 relaxation rates measurement

The two-dimensional $^{1}H^{15}N$ HSQC pulse sequence, for measuring heteroatom T1 relaxation time[137,138] (Fig. II.16), starts with a pre-scan period, usually defined by a long recycle delay, d1, to allow the recovery of the ^{1}H magnetization to a pre-equilibrium state just before to start the sequence.

After the initial ^{1}H excitation, by applying a ^{1}H 90 degree pulse, a fixed period, d26, optimized to 1/4J(NH) is applied to allow evolution of heteronuclear J(NH) coupling costant. Simoultaneous ^{1}H and ^{15}N 180 degree pulse is applied to refocus chemical shift evolution. At the end of the fixed period, d26, antiphase ^{1}H magnetization, with respect to ^{15}N, is obtained. A double 90 degree pulse both to ^{1}H and ^{15}N is applied to transfer magnetization from ^{1}H to ^{15}N. The above sequence described is known as Insensitive Nuclei Enhancement by Polarization Transfer (INEPT). Two refocused INEPT-type sequences are employed to transfer magnetization from the directly bound protons to the heteronucleus and back to protons for detection.

A water-selective 90º pulse creates transverse magnetization for the water whereas other signals are unaffected. The watergate block is usually applied during the retro-INEPT block of the standard HSQC pulse sequence in order to improve solvent

suppression in aqueous samples. The selective 90 degree pulse applied on the water resonance is defined by the pulse lenght for 90 degree shaped pulse (p11) and the power level for 90 degree shaped pulse (sp1). The gradients G1, G2, G3 and G4 dephases water signal. A way to suppress the protein resonances is to place two orthogonal composite-pulse spin lock trains (e.g., DIPSI-2) at the beginning of the relaxation period of T1 pulse sequence. ^{15}N T1 values were measured from the spectra recorded with three different durations of the delay T: T = 0.5, 1.0 and 1.5 seconds.

During the evolution time, t1, the ^{15}N chemical shift and ^{1}H-^{15}N coupling evolve. The ^{1}H 180° pulse in the center of the evolution time refocuses the coupling and as a result decouples protons in the F1 (^{15}N) domain of the spectrum. During the retro-INEPT sequence, optimized to 1/4J(NH), ^{15}N magnetizzation is transferred back to ^{1}H nuclei by applying a 90 degree pulse on ^{15}N. Simoultaneous 180 degree pulse, applied to ^{1}H and ^{15}N, refocus the ^{1}H chemical shift evolution. The pulse sequence-end was achieved by Globally optimized Alternating-phase Rectangular Pulses (GARP) to suppress ^{15}N-^{1}H heteronuclear scalar coupling during acquisition.

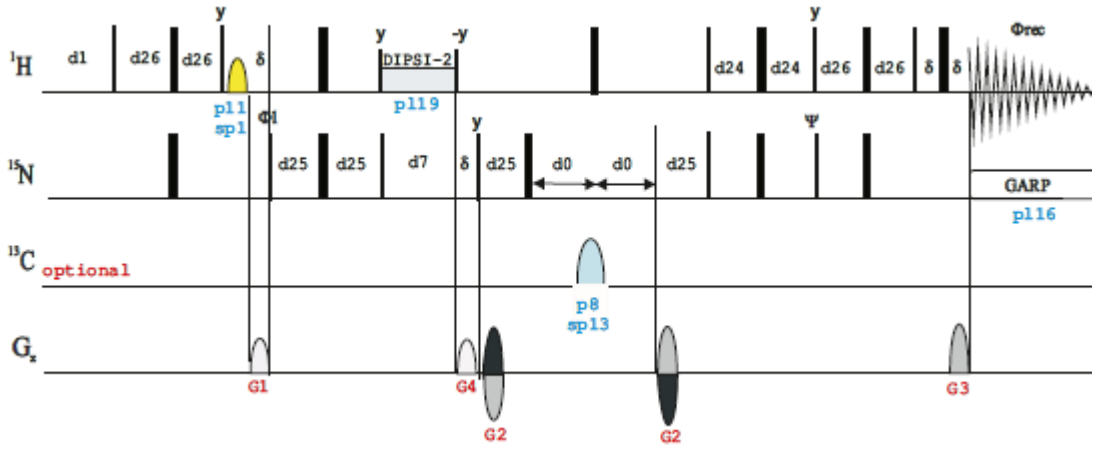


Figure II.16. Pulse sequence of longitudinal relaxation time T1 (hsqct1etf3gpsi3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

T2 relaxation rates measurement

The two-dimensional ^{1}H^{15}N HSQC pulse sequence for measuring heteroatom T2 relaxation time[139] (Fig. II.17) starts with a pre-scan period, usually defined by a long recycle delay, d1, to allow the recovery of the ^{1}H magnetization to a pre-equilibrium state just before to start the sequence. After the initial ^{1}H excitation, by applying a ^{1}H 90 degree

pulse, a fixed period, d26, optimized to 1/4J(NH) is applied to allow evolution of heteronuclear J(NH) coupling costant. Simoultaneous ^{1}H and ^{15}N 180 degree pulse is applied to refocus chemical shift evolution. At the end of the fixed period, d26, antiphase ^{1}H magnetization, with respect to ^{15}N, is obtained. A double 90 degree pulse both to ^{1}H and ^{15}N is applied to transfer magnetization from ^{1}H to ^{15}N. Two refocused INEPT-type sequences are employed to transfer magnetization from the directly bound protons to the heteronucleus and back to protons for detection. The sequence proceeds at the same way as T1 pulse sequence except for the relax bloch, where several 180° pulse are applied in ^{15}N to remove the inhomogeneities of the magnetic field. ^{15}N T2 values were measured from the spectra recorded with five different durations of the delay T: T = 0.05, 0.1, 0.150, 0.2 and 0.250 seconds. During the evolution time, t1, the ^{15}N chemical shift and ^{1}H-^{15}N coupling evolves. In order to measure T2 of nitrogen, spectra were recorded with the application of a 180 degree pulse on proton inthe middle of t1 evolution period.

During the retro-INEPT sequence, optimized to 1/4J(NH), ^{15}N magnetizzation is transferred back to ^{1}H nuclei by applying a 90 degree pulse on ^{15}N. Simoultaneous 180 degree pulse, applied to ^{1}H and ^{15}N, refocus the ^{1}H chemical shift evolution. The pulse sequence-end was achieved by Globally optimized Alternating-phase Rectangular Pulses (GARP) to suppress ^{15}N-^{1}H heteronuclear scalar coupling during acquisition.

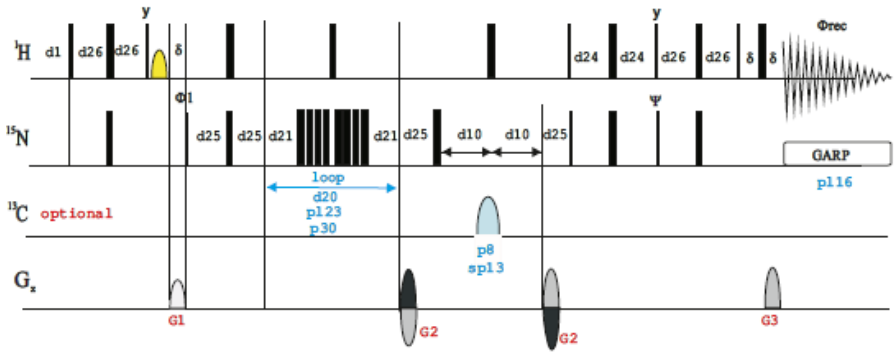


Figure II.17. Pulse sequence of transverse relaxation time T2 (hsqct2etf3gpsi3d). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

H-N NOE measurement

For heteronuclear NOE pulse sequence[140] (Fig. II.18), an INEPT sequence is used, where spectra are recorded in the presence and absence of ^{1}H saturation during the relaxation T

delay prior to the starting of 90 degree pulse on ^{15}N. The application of outdistanced 120 degree pulses, on proton, leads to its saturation.

Heteronuclear NOE calculation is measured by longitudinal magnetization of nitrogen and thus magnetization, in the pulse sequence, have to start on the nitrogen spin since the heteronuclear NOE is calculated from its longitudinal magnetization.

The selective 90 degree pulse applied on the water resonance is defined by the pulse lenght for 90 degree shaped pulse (p11) and the power level for 90 degree shaped pulse (sp1). The gradients G1, G2 and G3 dephases water signal. During the evolution time, t1, the ^{15}N chemical shift and ^{1}H-^{15}N coupling evolve. The ^{1}H 180° pulse in the center of the evolution time refocuses the coupling and as a result decouples protons in the F1 (^{15}N) domain of the spectrum. During the retro-INEPT sequence, optimized to 1/4J(NH), ^{15}N magnetizzation is transferred to ^{1}H nuclei by applying a 90 degree pulse on ^{15}N. Simoultaneous 180 degree pulse, applied to ^{1}H and ^{15}N, refocus the ^{1}H chemical shift evolution. Magnetization is subsequently transferred via a refocused INEPT sequence to the directly coupled proton for the detection.

The pulse sequence-end was achieved by Globally optimized Alternating-phase Rectangular Pulses (GARP) to suppress ^{15}N-^{1}H heteronuclear scalar coupling during acquisition.

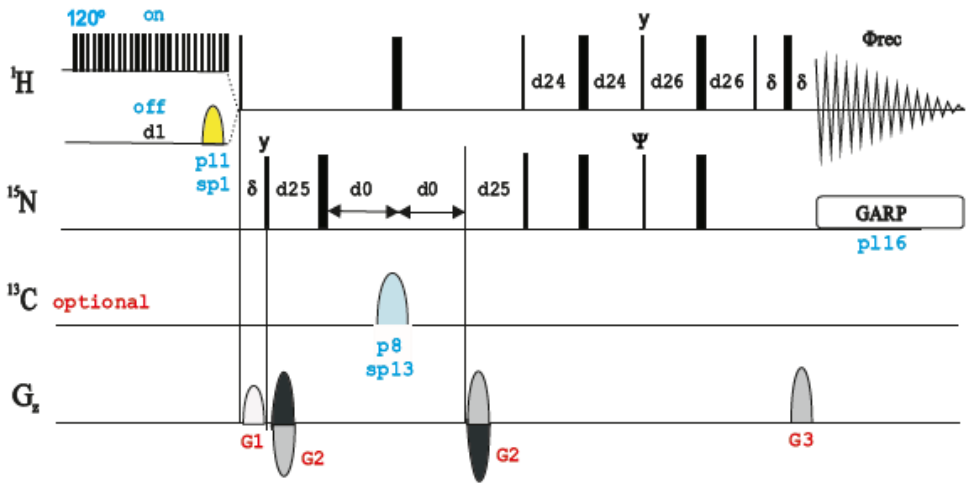


Figure II.18. Pulse sequence of heteronuclear ^{15}N-^{1}H NOEs (hsqcnoef3gpsi). In all sequence narrow and wide rectangles represent 90° and 180° pulses respectively. Pulses and gradients are indicated by semi-ellipses

ANNEX III.1: Expasy protein parameters

Construct 1: HisTag-Thioredoxin-TEV cleavage site -Pvfp5β

```
        10         20         30         40         50         60
MKSSHHHHHH GSSMSDKIIH LTDDSFDTDV LKADGAILVD FWAEWCGPCK MIAPILDEIA

        70         80         90        100        110        120
DEYQGKLTVA KLNIDQNPGT APKYGIRGIP TLLLFKNGEV AATKVGALSK GQLKEFLDAN

       130        140        150        160        170        180
LAGIEENLYF QGVYYPNPCS PYPCRNGGTC KKRGLYSYKC YCRKGYTGKN CQYNACFPNP

       190        200        210
CLNGGTCGYV YGYPYYKCSC PYGYYGKQCQ LKKY
```

Construct 2: HisTag-TEV cleavage site -Pvfp5β

```
        10         20         30         40         50         60
MKSSHHHHHH GSSMENLYFQ GVYYPNPCSP YPCRNGGTCK KRGLYSYKCY CRKGYTGKNC

        70         80         90        100
QYNACFPNPC LNGGTCGYVY GYPYYKCSCP YGYYGKQCQLKKY
```

	HisTag-Trx-TEVCleavageSite-Pvfp-5β		HisTag-TEVCleavageSite-Pvfp-5β		After TEV Cleavage Pvfp-5β	
Number of amino acids	214		103		83	
Molecular weight	23873.30		11916.55		9502.92 Da	
Theoretical pI	8.55		9.19		9.21	
Amino acid composition	Ala (A)	13 6.1%	Ala (A)	1 1.0%	Ala (A)	1 1.2%
	Arg (R)	4 1.9%	Arg (R)	3 2.9%	Arg (R)	3 3.6%
	Asn (N)	11 5.1%	Asn (N)	7 6.8%	Asn (N)	6 7.2%
	Asp (D)	11 5.1%	Asp (D)	0 0.0%	Asp (D)	0 0.0%
	Cys (C)	14 6.5%	Cys (C)	12 11.7%	Cys (C)	12 14.5%
	Gln (Q)	7 3.3%	Gln (Q)	4 3.9%	Gln (Q)	3 3.6%
	Glu (E)	7 3.3%	Glu (E)	1 1.0%	Glu (E)	0 0.0%
	Gly (G)	23 10.7%	Gly (G)	13 12.6%	Gly (G)	12 14.5%
	His (H)	7 3.3%	His (H)	6 5.8%	His (H)	0 0.0%
	Ile (I)	10 4.7%	Ile (I)	0 0.0%	Ile (I)	0 0.0%
	Leu (L)	17 7.9%	Leu (L)	4 3.9%	Leu (L)	3 3.6%
	Lys (K)	20 9.3%	Lys (K)	10 9.7%	Lys (K)	9 10.8%
	Met (M)	3 1.4%	Met (M)	2 1.9%	Met (M)	0 0.0%
	Phe (F)	6 2.8%	Phe (F)	2 1.9%	Phe (F)	1 1.2%
	Pro (P)	13 6.1%	Pro (P)	8 7.8%	Pro (P)	8 9.6%
	Ser (S)	10 4.7%	Ser (S)	7 6.8%	Ser (S)	3 3.6%
	Thr (T)	9 4.2%	Thr (T)	3 2.9%	Thr (T)	3 3.6%
	Trp (W)	2 0.9%	Trp (W)	0 0.0%	Trp (W)	0 0.0%
	Tyr (Y)	20 9.3%	Tyr (Y)	18 17.5%	Tyr (Y)	17 20.5%
	Val (V)	7 3.3%	Val (V)	2 1.9%	Val (V)	2 2.4%
	Pyl (O)	0 0.0%	Pyl (O)	0 0.0%	Pyl (O)	0 0.0%
	Sec (U)	0 0.0%	Sec (U)	0 0.0%	Sec (U)	0 0.0%

Total number of negatively Charged (Asp + Glu)	18	1	0
Total number of positively charged(Arg + Lys)	24	13	12
Ext. coefficient at 280 nm (assuming all pairs of Cys residues form cysteines)	41675	27570	26080
Abs 0.1% (=1 g/l) (assuming all pairs of Cys residues form cysteines)	1.746	2.314	2.744
Ext. coefficient at 280 nm (assuming all Cys residues are reduced)	40800	26820	25330
Abs 0.1% (=1 g/l) (assuming all Cys residues are reduced)	1.709	2.251	2.665

ANNEX III.2: Sanger DNA sequencing results for pHT-Pvfp-5β

Nucleotides sequence:

CNNNNNNNNNNNNNANNNNNNNGGNCNNNCTTTCGACTGAGCNCTTTCGTTTTATTTG
AAGCTTATTTAATTACCTGCAGGGAATTCTTAATACTTTTTCAGTTGGCATTGTTTGCC
GTAATAACCGTATGGGCAGCTACACTTGTAATACGGATAACCGTAAACGTAACCGCAAG
TGCCACCGTTCAGACACGGGTTCGGAAAACATGCGTTGTATTGACAGTTCTTGCCGGTA
TAACCTTTACGGCAATAACATTTATAAGAGTACAGACCACGTTTCTTACAGGTACCACC
GTTGCGGCACGGGTATGGAGAACACGGGTTCGGATAGTACACGCCTTGAAAGTACAGGT
TTTCCATAGAAGAACCATGGTGATGGTGATGGTGAGAAGATTTCATAATCTATGGTCCT
TGTTGGTCAATTGCTCGTGAAAACACCTAAACGGACTGGCTGTTTCCTGTGTGAAATTG
TTATCCGCTCACAATTCCACACATTATACGAGCCGATGATTAATTGTCAACAGCTCATT
TCAGAATATTTGCCAGAACCGTTATGATGTCGGCGCAAAAAACATTATCCAGAACGGGA
GTGCGCCTTGAGCGACACGAATTATGCAGTGATTTACGACCTGCACAGCCATACCACAG
CTTCCGATGGCTGCCTGACGCCAGAAGCATTGGTGCACCGTGCAGTCGATGATAAGCTG
TCAAACATGAGAATTGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCA
CTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTGTGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAG
ACGGGCAACAGCTGATTGCCCTTCACCGCNTGGCCCTGAGAGAGTTGCAGCAAGCGGTC
CACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGNTAACGGCGGGNTAT
ANNTGAGCTNNNTTCGGNATCGTCGTANNNACTACCGAAANATCCGNACCNNCGCGCNN
NCNGNCTCGGTAATGGNNNNNCATTGCNCCCNNCNNCATNTGATCGTTNGNNNNAGCAT
NGCAGNGGANNATNCCNTNANTNNGNATTNNATNGGNNNNTGAAANCNGNNNNTGGNNC
TNCAGNNCGNCNTTCCNTNNCNNN

Translated amino acid sequence:

6xHis – TEV cleavage site – Pvfp-5β protein

M K S S H H H H H H G S S **M** E N L Y F Q G V Y Y P N P C S P
Y P C R N G G T C K K R G L Y S Y K C Y C R K G Y T G K N C
Q Y N A C F P N P C L N G G T C G Y V Y G Y P Y Y K C S C P
Y G Y Y G K Q C Q L K K Y **Stop Codon**

ANNEX III.3: Backbone and side chains NMR resonances assignment

The columns of the following table of NMR resonances are:

- Column 1: Atom shift assign ID
- Column 2: Residue sequential number
- Column 3: Residue type
- Column 4: Atom name
- Column 5: Atom type
- Column 6: Chemical shift value
- Column 7: Chemical shift value error
- Column 8: Chemical shift ambiguity code

Chemical shift ambiguity code:

- 1: Unique
- 2: Ambiguity of geminal atoms or geminal methyl
- 3: Aromatic atoms on opposite sides of the ring (e.g. Tyr HE1 and HE2 protons)
- 4: Intraresidue ambiguities (e.g. Lys HG and HD protons)
- 5: Interresidue ambiguities (Lys 12 vs. Lys 27)
- 9: Ambiguous, specific ambiguity not defined

1	1	Gly	C	C	169.755	0.000	1
2	1	Gly	CA	C	43.370	0.089	1
3	2	Val	C	C	175.103	0.000	1
4	2	Val	CA	C	62.319	0.070	1
5	2	Val	CB	C	32.982	0.125	1
6	2	Val	CG1	C	20.553	0.188	1
7	2	Val	CG2	C	21.013	0.030	1
8	2	Val	H	H	8.347	0.002	1
9	2	Val	HA	H	4.050	0.005	1
10	2	Val	HB	H	1.903	0.008	1
11	2	Val	HG11	H	0.838	0.004	1
12	2	Val	HG12	H	0.838	0.004	1
13	2	Val	HG13	H	0.838	0.004	1
14	2	Val	HG21	H	0.718	0.015	1
15	2	Val	HG22	H	0.718	0.015	1
16	2	Val	HG23	H	0.718	0.015	1
17	2	Val	N	N	119.556	0.114	1
18	3	Tyr	C	C	174.324	0.000	1
19	3	Tyr	CA	C	57.592	0.092	1
20	3	Tyr	CB	C	39.268	0.077	1
21	3	Tyr	CD1	C	133.002	0.000	1
22	3	Tyr	CD2	C	133.002	0.000	1
23	3	Tyr	CE1	C	118.275	0.000	1
24	3	Tyr	CE2	C	118.275	0.000	1
25	3	Tyr	H	H	8.257	0.007	1

26	3	Tyr	HA	H	4.536	0.005	1
27	3	Tyr	HB2	H	2.840	0.033	2
28	3	Tyr	HB3	H	2.891	0.039	2
29	3	Tyr	HD1	H	7.058	0.011	1
30	3	Tyr	HD2	H	7.058	0.011	1
31	3	Tyr	HE1	H	6.781	0.007	1
32	3	Tyr	HE2	H	6.781	0.007	1
33	3	Tyr	N	N	124.949	0.099	1
34	4	Tyr	CA	C	55.194	0.033	1
35	4	Tyr	CB	C	39.020	0.092	1
36	4	Tyr	CD1	C	133.449	0.055	1
37	4	Tyr	CD2	C	133.449	0.055	1
38	4	Tyr	CE1	C	118.126	0.000	1
39	4	Tyr	CE2	C	118.126	0.000	1
40	4	Tyr	H	H	7.985	0.008	1
41	4	Tyr	HA	H	4.726	0.016	1
42	4	Tyr	HB2	H	2.686	0.005	2
43	4	Tyr	HB3	H	2.950	0.010	2
44	4	Tyr	HD1	H	7.085	0.012	1
45	4	Tyr	HD2	H	7.085	0.012	1
46	4	Tyr	HE1	H	6.768	0.009	1
47	4	Tyr	HE2	H	6.768	0.009	1
48	4	Tyr	N	N	125.511	0.273	1
49	5	Pro	C	C	175.580	0.000	1
50	5	Pro	CA	C	62.973	0.036	1
51	5	Pro	CB	C	32.286	0.060	1
52	5	Pro	CD	C	50.704	0.050	1
53	5	Pro	CG	C	27.368	0.047	1
54	5	Pro	HA	H	4.268	0.005	1
55	5	Pro	HB2	H	1.927	0.033	2
56	5	Pro	HB3	H	2.284	0.004	2
57	5	Pro	HD2	H	3.513	0.023	2
58	5	Pro	HD3	H	3.584	0.006	2
59	5	Pro	HG2	H	2.286	0.003	1
60	5	Pro	HG3	H	2.285	0.003	1
61	6	Asn	CA	C	49.914	0.055	1
62	6	Asn	CB	C	39.445	0.071	1
63	6	Asn	H	H	8.424	0.005	1
64	6	Asn	HA	H	4.986	0.006	1
65	6	Asn	HB2	H	2.826	0.014	2
66	6	Asn	HB3	H	3.075	0.007	2
67	6	Asn	N	N	118.622	0.190	1
68	7	Pro	C	C	176.184	0.000	1
69	7	Pro	CA	C	63.904	0.035	1
70	7	Pro	CB	C	32.892	0.105	1
71	7	Pro	CD	C	50.645	0.061	1
72	7	Pro	CG	C	27.633	0.094	1
73	7	Pro	HA	H	4.515	0.007	1
74	7	Pro	HB2	H	1.818	0.008	2
75	7	Pro	HB3	H	2.113	0.005	2
76	7	Pro	HD2	H	3.900	0.004	2
77	7	Pro	HD3	H	3.739	0.008	2
78	7	Pro	HG2	H	1.857	0.004	2

79	7	Pro	HG3	H	2.186	0.008	2
80	8	Cys	C	C	173.096	0.000	1
81	8	Cys	CA	C	53.836	0.063	1
82	8	Cys	CB	C	41.206	0.085	1
83	8	Cys	H	H	7.855	0.006	1
84	8	Cys	HA	H	4.256	0.010	1
85	8	Cys	HB2	H	3.317	0.018	2
86	8	Cys	HB3	H	2.789	0.017	2
87	8	Cys	N	N	112.389	0.068	1
88	9	Ser	CA	C	55.042	0.091	1
89	9	Ser	CB	C	65.045	0.082	1
90	9	Ser	H	H	7.464	0.005	1
91	9	Ser	HA	H	4.527	0.004	1
92	9	Ser	HB2	H	3.787	0.009	2
93	9	Ser	HB3	H	3.520	0.004	2
94	9	Ser	N	N	116.308	0.059	1
95	10	Pro	C	C	176.412	0.000	1
96	10	Pro	CA	C	62.710	0.072	1
97	10	Pro	CB	C	34.159	0.064	1
98	10	Pro	CD	C	50.391	0.085	1
99	10	Pro	CG	C	25.565	0.070	1
100	10	Pro	HA	H	4.652	0.003	1
101	10	Pro	HB2	H	2.390	0.007	2
102	10	Pro	HB3	H	1.882	0.030	2
103	10	Pro	HD2	H	3.526	0.003	1
104	10	Pro	HD3	H	3.526	0.003	1
105	10	Pro	HG2	H	1.851	0.022	2
106	10	Pro	HG3	H	1.820	0.027	2
107	11	Tyr	CA	C	57.667	0.200	1
108	11	Tyr	CB	C	39.414	0.054	1
109	11	Tyr	CD1	C	133.331	0.011	1
110	11	Tyr	CD2	C	133.331	0.011	1
111	11	Tyr	CE1	C	118.909	0.020	1
112	11	Tyr	CE2	C	118.909	0.020	1
113	11	Tyr	H	H	8.868	0.009	1
114	11	Tyr	HA	H	4.456	0.010	1
115	11	Tyr	HB2	H	3.051	0.023	2
116	11	Tyr	HB3	H	3.019	0.029	2
117	11	Tyr	HD1	H	7.284	0.014	1
118	11	Tyr	HD2	H	7.284	0.014	1
119	11	Tyr	HE1	H	7.065	0.006	1
120	11	Tyr	HE2	H	7.065	0.006	1
121	11	Tyr	N	N	126.888	0.057	1
122	13	Cys	C	C	173.417	0.000	1
123	13	Cys	CA	C	52.859	0.064	1
124	13	Cys	CB	C	38.192	0.055	1
125	13	Cys	HA	H	4.460	0.003	1
126	13	Cys	HB2	H	2.483	0.008	1
127	13	Cys	HB3	H	2.483	0.008	1
128	14	Arg	C	C	176.423	0.000	1
129	14	Arg	CA	C	54.443	0.115	1
130	14	Arg	CB	C	32.143	0.109	1
131	14	Arg	CD	C	43.134	0.086	1

132	14	Arg	CG	C	27.133	0.097	1
133	14	Arg	H	H	8.604	0.006	1
134	14	Arg	HA	H	4.403	0.015	1
135	14	Arg	HB2	H	1.868	0.027	1
136	14	Arg	HB3	H	1.833	0.000	2
137	14	Arg	HD2	H	3.090	0.005	1
138	14	Arg	HD3	H	3.090	0.005	1
139	14	Arg	HG2	H	1.657	0.008	1
140	14	Arg	HG3	H	1.657	0.008	1
141	14	Arg	N	N	122.996	0.161	1
142	15	Asn	C	C	176.705	0.000	1
143	15	Asn	CA	C	54.879	0.038	1
144	15	Asn	CB	C	37.208	0.201	1
145	15	Asn	H	H	9.168	0.004	1
146	15	Asn	HA	H	4.125	0.009	1
147	15	Asn	HB2	H	2.073	0.010	2
148	15	Asn	HB3	H	1.516	0.017	2
149	15	Asn	N	N	114.218	0.084	1
150	16	Gly	C	C	175.534	0.000	1
151	16	Gly	CA	C	45.869	0.102	1
152	16	Gly	H	H	8.858	0.006	1
153	16	Gly	HA2	H	3.733	0.020	2
154	16	Gly	HA3	H	3.907	0.011	2
155	16	Gly	N	N	107.200	0.081	1
156	17	Gly	C	C	172.555	0.000	1
157	17	Gly	CA	C	45.768	0.102	1
158	17	Gly	H	H	7.451	0.008	1
159	17	Gly	HA2	H	3.505	0.012	2
160	17	Gly	HA3	H	4.445	0.013	2
161	17	Gly	N	N	105.927	0.116	1
162	18	Thr	C	C	172.767	0.000	1
163	18	Thr	CA	C	62.850	0.054	1
164	18	Thr	CB	C	71.161	0.084	1
165	18	Thr	CG2	C	21.880	0.061	1
166	18	Thr	H	H	8.934	0.009	1
167	18	Thr	HA	H	4.556	0.006	1
168	18	Thr	HB	H	4.210	0.009	1
169	18	Thr	HG21	H	1.275	0.007	1
170	18	Thr	HG22	H	1.275	0.007	1
171	18	Thr	HG23	H	1.275	0.007	1
172	18	Thr	N	N	119.377	0.091	1
173	19	Cys	C	C	173.268	0.000	1
174	19	Cys	CA	C	57.323	0.065	1
175	19	Cys	CB	C	43.114	0.197	1
176	19	Cys	H	H	8.517	0.007	1
177	19	Cys	HA	H	4.799	0.016	1
178	19	Cys	HB2	H	2.841	0.007	2
179	19	Cys	HB3	H	2.285	0.005	2
180	19	Cys	N	N	127.982	0.119	1
181	20	Lys	C	C	175.297	0.000	1
182	20	Lys	CA	C	54.707	0.157	1
183	20	Lys	CB	C	35.628	0.063	1
184	20	Lys	CD	C	28.999	0.091	1

185	20	Lys	CE	C	42.243	0.027	1
186	20	Lys	CG	C	25.056	0.044	1
187	20	Lys	H	H	9.333	0.007	1
188	20	Lys	HA	H	4.876	0.012	1
189	20	Lys	HB2	H	1.994	0.008	2
190	20	Lys	HB3	H	1.857	0.011	2
191	20	Lys	HD2	H	1.770	0.005	2
192	20	Lys	HD3	H	1.645	0.008	2
193	20	Lys	HE2	H	2.835	0.019	2
194	20	Lys	HE3	H	2.778	0.005	2
195	20	Lys	HG2	H	1.470	0.007	1
196	20	Lys	HG3	H	1.470	0.007	1
197	20	Lys	N	N	130.399	0.555	1
198	21	Lys	C	C	176.132	0.000	1
199	21	Lys	CA	C	57.789	0.065	1
200	21	Lys	CB	C	32.880	0.120	1
201	21	Lys	CD	C	29.702	0.089	1
202	21	Lys	CE	C	41.887	0.177	1
203	21	Lys	CG	C	25.333	0.033	1
204	21	Lys	H	H	8.840	0.010	1
205	21	Lys	HA	H	4.238	0.008	1
206	21	Lys	HB2	H	1.911	0.016	2
207	21	Lys	HB3	H	1.897	0.007	2
208	21	Lys	HD2	H	1.715	0.021	2
209	21	Lys	HD3	H	1.693	0.007	2
210	21	Lys	HE2	H	2.986	0.010	2
211	21	Lys	HE3	H	2.976	0.009	2
212	21	Lys	HG2	H	1.563	0.010	2
213	21	Lys	HG3	H	1.349	0.009	2
214	21	Lys	N	N	124.245	0.104	1
215	22	Arg	C	C	174.902	0.000	1
216	22	Arg	CA	C	54.565	0.035	1
217	22	Arg	CB	C	31.234	0.085	1
218	22	Arg	CD	C	43.464	0.056	1
219	22	Arg	CG	C	26.701	0.043	1
220	22	Arg	H	H	8.338	0.006	1
221	22	Arg	HA	H	4.569	0.005	1
222	22	Arg	HB2	H	1.644	0.005	2
223	22	Arg	HB3	H	1.311	0.006	2
224	22	Arg	HD2	H	3.113	0.005	1
225	22	Arg	HD3	H	3.113	0.005	1
226	22	Arg	HG2	H	1.500	0.011	2
227	22	Arg	HG3	H	1.439	0.009	2
228	22	Arg	N	N	126.068	0.132	1
229	23	Gly	C	C	173.115	0.000	1
230	23	Gly	CA	C	44.983	0.089	1
231	23	Gly	H	H	8.235	0.003	1
232	23	Gly	HA2	H	3.740	0.011	2
233	23	Gly	HA3	H	4.037	0.010	2
234	23	Gly	N	N	111.621	0.102	1
235	24	Leu	C	C	178.074	0.000	1
236	24	Leu	CA	C	56.922	0.108	1
237	24	Leu	CB	C	42.095	0.054	1

238	24	Leu	CD1	C	26.896	0.062	1
239	24	Leu	CD2	C	26.900	0.058	1
240	24	Leu	H	H	8.246	0.006	1
241	24	Leu	HA	H	3.925	0.008	1
242	24	Leu	HB2	H	0.959	0.016	2
243	24	Leu	HB3	H	0.698	0.012	2
244	24	Leu	HD11	H	1.387	0.004	1
245	24	Leu	HD12	H	1.387	0.004	1
246	24	Leu	HD13	H	1.387	0.004	1
247	24	Leu	HD21	H	1.387	0.004	1
248	24	Leu	HD22	H	1.387	0.004	1
249	24	Leu	HD23	H	1.387	0.004	1
250	24	Leu	HG	H	2.140	0.022	1
251	24	Leu	N	N	119.840	0.100	1
252	25	Tyr	C	C	175.720	0.000	1
253	25	Tyr	CA	C	56.882	0.078	1
254	25	Tyr	CB	C	39.927	0.157	1
255	25	Tyr	CD1	C	133.626	0.041	1
256	25	Tyr	CD2	C	133.626	0.041	1
257	25	Tyr	CE1	C	118.108	0.032	1
258	25	Tyr	CE2	C	118.108	0.032	1
259	25	Tyr	H	H	8.103	0.005	1
260	25	Tyr	HA	H	4.860	0.017	1
261	25	Tyr	HB2	H	3.378	0.011	2
262	25	Tyr	HB3	H	2.762	0.009	2
263	25	Tyr	HD1	H	7.105	0.014	1
264	25	Tyr	HD2	H	7.105	0.014	1
265	25	Tyr	HE1	H	6.801	0.009	1
266	25	Tyr	HE2	H	6.801	0.009	1
267	25	Tyr	N	N	111.822	0.068	1
268	26	Ser	C	C	171.746	0.000	1
269	26	Ser	CA	C	56.762	0.099	1
270	26	Ser	CB	C	64.911	0.064	1
271	26	Ser	H	H	7.277	0.006	1
272	26	Ser	HA	H	4.921	0.009	1
273	26	Ser	HB2	H	3.789	0.007	1
274	26	Ser	HB3	H	3.789	0.007	1
275	26	Ser	N	N	113.846	0.067	1
276	27	Tyr	C	C	173.252	0.000	1
277	27	Tyr	CA	C	55.828	0.085	1
278	27	Tyr	CB	C	42.182	0.084	1
279	27	Tyr	CD1	C	133.943	0.035	1
280	27	Tyr	CD2	C	133.943	0.035	1
281	27	Tyr	CE1	C	118.109	0.021	1
282	27	Tyr	CE2	C	118.109	0.021	1
283	27	Tyr	H	H	8.022	0.004	1
284	27	Tyr	HA	H	5.343	0.007	1
285	27	Tyr	HB2	H	2.978	0.020	2
286	27	Tyr	HB3	H	2.491	0.009	2
287	27	Tyr	HD1	H	6.755	0.008	1
288	27	Tyr	HD2	H	6.755	0.008	1
289	27	Tyr	HE1	H	6.806	0.013	1
290	27	Tyr	HE2	H	6.806	0.013	1

291	27	Tyr	N	N	117.714	0.096	1
292	28	Lys	C	C	175.007	0.000	1
293	28	Lys	CA	C	55.024	0.074	1
294	28	Lys	CB	C	36.823	0.113	1
295	28	Lys	CD	C	29.379	0.162	1
296	28	Lys	CE	C	42.218	0.027	1
297	28	Lys	H	H	9.156	0.005	1
298	28	Lys	HA	H	4.440	0.007	1
299	28	Lys	HB2	H	1.612	0.007	2
300	28	Lys	HB3	H	1.539	0.013	2
301	28	Lys	HD2	H	1.553	0.020	1
302	28	Lys	HD3	H	1.553	0.020	1
303	28	Lys	HE2	H	2.595	0.013	2
304	28	Lys	HE3	H	2.545	0.007	2
305	28	Lys	N	N	121.083	0.136	1
306	29	Cys	C	C	173.711	0.000	1
307	29	Cys	CA	C	54.504	0.090	1
308	29	Cys	CB	C	42.709	0.113	1
309	29	Cys	H	H	9.140	0.007	1
310	29	Cys	HA	H	5.386	0.013	1
311	29	Cys	HB2	H	3.122	0.022	2
312	29	Cys	HB3	H	2.723	0.007	2
313	29	Cys	N	N	121.822	0.083	1
314	30	Tyr	C	C	176.463	0.000	1
315	30	Tyr	CA	C	57.518	0.117	1
316	30	Tyr	CB	C	37.636	0.075	1
317	30	Tyr	CD1	C	133.206	0.031	1
318	30	Tyr	CD2	C	133.206	0.031	1
319	30	Tyr	CE1	C	118.133	0.094	1
320	30	Tyr	CE2	C	118.133	0.094	1
321	30	Tyr	H	H	9.450	0.003	1
322	30	Tyr	HA	H	4.571	0.004	1
323	30	Tyr	HB2	H	3.128	0.018	2
324	30	Tyr	HB3	H	2.939	0.008	2
325	30	Tyr	HD1	H	7.215	0.008	1
326	30	Tyr	HD2	H	7.215	0.008	1
327	30	Tyr	HE1	H	6.769	0.009	1
328	30	Tyr	HE2	H	6.769	0.009	1
329	30	Tyr	N	N	125.226	0.101	1
330	31	Cys	C	C	175.920	0.000	1
331	31	Cys	CA	C	53.385	0.040	1
332	31	Cys	CB	C	38.051	0.117	1
333	31	Cys	H	H	8.878	0.006	1
334	31	Cys	HA	H	4.764	0.022	1
335	31	Cys	HB2	H	3.313	0.016	2
336	31	Cys	HB3	H	2.688	0.013	2
337	31	Cys	N	N	123.825	0.097	1
338	32	Arg	C	C	175.432	0.000	1
339	32	Arg	CA	C	56.157	0.077	1
340	32	Arg	CB	C	31.292	0.112	1
341	32	Arg	CD	C	43.680	0.034	1
342	32	Arg	CG	C	27.996	0.084	1
343	32	Arg	H	H	8.511	0.007	1

344	32	Arg	HA	H	4.433	0.009	1
345	32	Arg	HB2	H	2.101	0.007	2
346	32	Arg	HB3	H	1.678	0.007	2
347	32	Arg	HD2	H	3.286	0.002	1
348	32	Arg	HD3	H	3.286	0.002	1
349	32	Arg	HG2	H	1.680	0.007	1
350	32	Arg	HG3	H	1.680	0.007	1
351	32	Arg	N	N	121.205	0.135	1
352	33	Lys	C	C	176.029	0.000	1
353	33	Lys	CA	C	58.431	0.091	1
354	33	Lys	CB	C	32.414	0.061	1
355	33	Lys	CD	C	29.325	0.083	1
356	33	Lys	CE	C	41.869	0.030	1
357	33	Lys	CG	C	24.659	0.045	1
358	33	Lys	H	H	8.371	0.008	1
359	33	Lys	HA	H	3.866	0.007	1
360	33	Lys	HB2	H	1.729	0.004	2
361	33	Lys	HB3	H	1.612	0.007	2
362	33	Lys	HD2	H	1.547	0.006	2
363	33	Lys	HD3	H	1.536	0.007	2
364	33	Lys	HE2	H	2.707	0.007	2
365	33	Lys	HE3	H	2.604	0.005	2
366	33	Lys	HG2	H	1.087	0.007	1
367	33	Lys	HG3	H	1.087	0.007	1
368	33	Lys	N	N	121.575	0.072	1
369	34	Gly	C	C	172.658	0.000	1
370	34	Gly	CA	C	44.493	0.076	1
371	34	Gly	H	H	8.404	0.015	1
372	34	Gly	HA2	H	2.991	0.007	2
373	34	Gly	HA3	H	3.405	0.009	2
374	34	Gly	N	N	112.720	0.064	1
375	35	Tyr	C	C	175.650	0.000	1
376	35	Tyr	CA	C	57.024	0.179	1
377	35	Tyr	CB	C	40.105	0.058	1
378	35	Tyr	CD1	C	134.062	0.000	1
379	35	Tyr	CD2	C	134.062	0.000	1
380	35	Tyr	CE1	C	117.720	0.051	1
381	35	Tyr	CE2	C	117.720	0.051	1
382	35	Tyr	H	H	7.763	0.005	1
383	35	Tyr	HA	H	5.561	0.006	1
384	35	Tyr	HB2	H	2.927	0.014	2
385	35	Tyr	HB3	H	2.899	0.007	2
386	35	Tyr	HD1	H	6.891	0.016	1
387	35	Tyr	HD2	H	6.891	0.016	1
388	35	Tyr	HE1	H	6.651	0.016	1
389	35	Tyr	HE2	H	6.651	0.016	1
390	35	Tyr	N	N	117.615	0.100	1
391	36	Thr	C	C	173.854	0.000	1
392	36	Thr	CA	C	60.094	0.059	1
393	36	Thr	CB	C	71.695	0.075	1
394	36	Thr	CG2	C	20.815	0.043	1
395	36	Thr	H	H	9.014	0.003	1
396	36	Thr	HA	H	4.719	0.022	1

397	36	Thr	HB	H	4.133	0.007	1
398	36	Thr	HG21	H	0.909	0.012	1
399	36	Thr	HG22	H	0.909	0.012	1
400	36	Thr	HG23	H	0.909	0.012	1
401	36	Thr	N	N	112.819	0.073	1
402	37	Gly	C	C	176.798	0.000	1
403	37	Gly	CA	C	43.826	0.183	1
404	37	Gly	H	H	8.128	0.006	1
405	37	Gly	HA2	H	4.817	0.025	2
406	37	Gly	HA3	H	4.013	0.013	2
407	37	Gly	N	N	107.007	0.086	1
408	38	Lys	C	C	176.480	0.000	1
409	38	Lys	CA	C	59.732	0.057	1
410	38	Lys	CB	C	33.296	0.089	1
411	38	Lys	CD	C	29.491	0.061	1
412	38	Lys	CE	C	42.313	0.071	1
413	38	Lys	CG	C	24.787	0.082	1
414	38	Lys	H	H	9.212	0.007	1
415	38	Lys	HA	H	3.996	0.010	1
416	38	Lys	HB2	H	1.998	0.009	2
417	38	Lys	HB3	H	1.722	0.011	2
418	38	Lys	HD2	H	1.688	0.009	1
419	38	Lys	HD3	H	1.688	0.009	1
420	38	Lys	HE2	H	2.908	0.007	2
421	38	Lys	HE3	H	2.867	0.004	2
422	38	Lys	HG2	H	1.448	0.015	2
423	38	Lys	HG3	H	1.424	0.016	2
424	38	Lys	N	N	124.454	0.153	1
425	39	Asn	C	C	175.495	0.000	1
426	39	Asn	CA	C	52.090	0.091	1
427	39	Asn	CB	C	39.692	0.090	1
428	39	Asn	H	H	8.462	0.004	1
429	39	Asn	HA	H	5.534	0.013	1
430	39	Asn	HB2	H	3.481	0.011	2
431	39	Asn	HB3	H	2.222	0.006	2
432	39	Asn	N	N	111.894	0.056	1
433	40	Cys	C	C	174.273	0.000	1
434	40	Cys	CA	C	54.935	0.076	1
435	40	Cys	CB	C	36.514	0.082	1
436	40	Cys	H	H	8.010	0.005	1
437	40	Cys	HA	H	4.093	0.013	1
438	40	Cys	HB2	H	3.519	0.011	2
439	40	Cys	HB3	H	3.181	0.005	2
440	40	Cys	N	N	113.752	0.075	1
441	41	Gln	C	C	175.564	0.000	1
442	41	Gln	CA	C	57.203	0.093	1
443	41	Gln	CB	C	28.071	0.165	1
444	41	Gln	H	H	9.728	0.002	1
445	41	Gln	HA	H	3.918	0.012	1
446	41	Gln	HB2	H	2.006	0.010	2
447	41	Gln	HB3	H	1.543	0.010	2
448	41	Gln	N	N	117.752	0.194	1
449	42	Tyr	C	C	174.081	0.000	1

450	42	Tyr	CA	C	56.568	0.066	1
451	42	Tyr	CB	C	38.483	0.124	1
452	42	Tyr	CD1	C	133.099	0.102	1
453	42	Tyr	CD2	C	133.099	0.102	1
454	42	Tyr	CE1	C	118.117	0.000	1
455	42	Tyr	CE2	C	118.117	0.000	1
456	42	Tyr	H	H	8.740	0.006	1
457	42	Tyr	HA	H	5.098	0.009	1
458	42	Tyr	HB2	H	3.038	0.013	2
459	42	Tyr	HB3	H	2.912	0.007	2
460	42	Tyr	HD1	H	7.058	0.012	1
461	42	Tyr	HD2	H	7.058	0.012	1
462	42	Tyr	HE1	H	6.793	0.006	1
463	42	Tyr	HE2	H	6.793	0.006	1
464	42	Tyr	N	N	123.835	0.082	1
465	43	Asn	C	C	176.174	0.000	1
466	43	Asn	CA	C	51.975	0.046	1
467	43	Asn	CB	C	40.182	0.157	1
468	43	Asn	H	H	8.939	0.008	1
469	43	Asn	HA	H	4.680	0.010	1
470	43	Asn	HB2	H	3.381	0.010	2
471	43	Asn	HB3	H	2.735	0.020	2
472	43	Asn	N	N	121.455	0.118	1
473	44	Ala	C	C	176.844	0.000	1
474	44	Ala	CA	C	54.645	0.064	1
475	44	Ala	CB	C	20.109	0.054	1
476	44	Ala	H	H	8.874	0.005	1
477	44	Ala	HA	H	4.167	0.007	1
478	44	Ala	HB1	H	1.602	0.007	1
479	44	Ala	HB2	H	1.602	0.007	1
480	44	Ala	HB3	H	1.602	0.007	1
481	44	Ala	N	N	127.071	0.094	1
482	45	Cys	C	C	173.873	0.000	1
483	45	Cys	CA	C	52.124	0.071	1
484	45	Cys	CB	C	39.270	0.068	1
485	45	Cys	H	H	8.060	0.006	1
486	45	Cys	HA	H	5.003	0.009	1
487	45	Cys	HB2	H	3.189	0.007	2
488	45	Cys	HB3	H	2.396	0.007	2
489	45	Cys	N	N	107.778	0.080	1
490	46	Phe	CA	C	56.203	0.068	1
491	46	Phe	CB	C	42.454	0.106	1
492	46	Phe	CD1	C	131.637	0.037	1
493	46	Phe	CD2	C	131.637	0.037	1
494	46	Phe	CE1	C	131.598	0.028	1
495	46	Phe	CE2	C	131.598	0.028	1
496	46	Phe	CZ	C	130.054	0.020	1
497	46	Phe	H	H	7.433	0.005	1
498	46	Phe	HA	H	4.549	0.005	1
499	46	Phe	HB2	H	2.948	0.007	2
500	46	Phe	HB3	H	2.769	0.010	2
501	46	Phe	HD1	H	7.256	0.012	1
502	46	Phe	HD2	H	7.256	0.012	1

503	46	Phe	HE1	H	7.384	0.012	1
504	46	Phe	HE2	H	7.384	0.012	1
505	46	Phe	HZ	H	7.322	0.018	1
506	46	Phe	N	N	121.678	0.241	1
507	47	Pro	C	C	175.493	0.000	1
508	47	Pro	CA	C	62.759	0.090	1
509	47	Pro	CB	C	34.018	0.081	1
510	47	Pro	CD	C	49.915	0.048	1
511	47	Pro	CG	C	25.638	0.090	1
512	47	Pro	HA	H	3.778	0.006	1
513	47	Pro	HB2	H	2.024	0.005	2
514	47	Pro	HB3	H	1.649	0.015	2
515	47	Pro	HD2	H	3.392	0.007	2
516	47	Pro	HD3	H	3.343	0.017	2
517	47	Pro	HG2	H	1.682	0.015	2
518	47	Pro	HG3	H	1.517	0.004	2
519	48	Asn	C	C	175.513	0.000	1
520	48	Asn	CA	C	51.341	0.032	1
521	48	Asn	CB	C	38.438	0.163	1
522	48	Asn	H	H	8.593	0.004	1
523	48	Asn	HA	H	4.351	0.006	1
524	48	Asn	HB2	H	2.904	0.007	2
525	48	Asn	HB3	H	2.830	0.014	2
526	48	Asn	N	N	122.370	0.209	1
527	49	Pro	C	C	176.002	0.000	1
528	49	Pro	CA	C	63.087	0.056	1
529	49	Pro	CB	C	32.392	0.078	1
530	49	Pro	CD	C	51.472	0.100	1
531	49	Pro	CG	C	25.388	0.117	1
532	49	Pro	HA	H	4.249	0.006	1
533	49	Pro	HB2	H	1.692	0.019	2
534	49	Pro	HB3	H	1.416	0.006	2
535	49	Pro	HD2	H	4.018	0.006	2
536	49	Pro	HD3	H	3.808	0.007	2
537	49	Pro	HG2	H	0.803	0.011	2
538	49	Pro	HG3	H	1.112	0.006	2
539	50	Cys	C	C	173.832	0.000	1
540	50	Cys	CA	C	53.313	0.107	1
541	50	Cys	CB	C	38.307	0.045	1
542	50	Cys	H	H	7.769	0.008	1
543	50	Cys	HA	H	4.215	0.006	1
544	50	Cys	HB2	H	2.743	0.007	2
545	50	Cys	HB3	H	2.300	0.007	2
546	50	Cys	N	N	118.908	0.085	1
547	51	Leu	C	C	177.121	0.000	1
548	51	Leu	CA	C	54.359	0.063	1
549	51	Leu	CB	C	43.731	0.055	1
550	51	Leu	CD1	C	22.320	0.050	1
551	51	Leu	CD2	C	25.476	0.027	1
552	51	Leu	CG	C	27.451	0.065	1
553	51	Leu	H	H	8.490	0.005	1
554	51	Leu	HA	H	4.225	0.006	1
555	51	Leu	HB2	H	1.355	0.006	2

556	51	Leu	HB3	H	1.564	0.008	2
557	51	Leu	HD11	H	0.838	0.003	2
558	51	Leu	HD12	H	0.838	0.003	2
559	51	Leu	HD13	H	0.838	0.003	2
560	51	Leu	HD21	H	0.738	0.007	2
561	51	Leu	HD22	H	0.738	0.007	2
562	51	Leu	HD23	H	0.738	0.007	2
563	51	Leu	HG	H	1.722	0.004	1
564	51	Leu	N	N	125.516	0.067	1
565	52	Asn	C	C	174.005	0.000	1
566	52	Asn	CA	C	55.085	0.085	1
567	52	Asn	CB	C	37.007	0.093	1
568	52	Asn	H	H	8.604	0.007	1
569	52	Asn	HA	H	3.950	0.010	1
570	52	Asn	HB2	H	1.514	0.005	2
571	52	Asn	HB3	H	2.085	0.008	2
572	52	Asn	N	N	110.884	0.063	1
573	53	Gly	C	C	176.220	0.000	1
574	53	Gly	CA	C	45.017	0.056	1
575	53	Gly	H	H	8.263	0.005	1
576	53	Gly	HA2	H	4.072	0.009	2
577	53	Gly	HA3	H	3.535	0.007	2
578	53	Gly	N	N	102.621	0.066	1
579	54	Gly	C	C	171.949	0.000	1
580	54	Gly	CA	C	46.221	0.074	1
581	54	Gly	H	H	7.757	0.008	1
582	54	Gly	HA2	H	3.507	0.008	2
583	54	Gly	HA3	H	4.426	0.009	2
584	54	Gly	N	N	107.098	0.109	1
585	55	Thr	C	C	172.591	0.000	1
586	55	Thr	CA	C	61.032	0.132	1
587	55	Thr	CB	C	70.683	0.048	1
588	55	Thr	CG2	C	21.752	0.083	1
589	55	Thr	H	H	8.610	0.007	1
590	55	Thr	HA	H	4.481	0.008	1
591	55	Thr	HB	H	4.066	0.007	1
592	55	Thr	HG21	H	1.185	0.012	1
593	55	Thr	HG22	H	1.185	0.012	1
594	55	Thr	HG23	H	1.185	0.012	1
595	55	Thr	N	N	115.994	0.100	1
596	56	Cys	C	C	173.911	0.000	1
597	56	Cys	CA	C	57.490	0.085	1
598	56	Cys	CB	C	38.244	0.105	1
599	56	Cys	H	H	9.101	0.004	1
600	56	Cys	HA	H	4.919	0.008	1
601	56	Cys	HB2	H	3.174	0.008	2
602	56	Cys	HB3	H	2.775	0.009	2
603	56	Cys	N	N	128.309	0.125	1
604	57	Gly	C	C	171.646	0.000	1
605	57	Gly	CA	C	43.942	0.067	1
606	57	Gly	H	H	8.877	0.008	1
607	57	Gly	HA2	H	4.370	0.007	2
608	57	Gly	HA3	H	3.641	0.039	2

609	57	Gly	N	N	118.261	0.089	1
610	58	Tyr	C	C	176.708	0.000	1
611	58	Tyr	CA	C	57.186	0.092	1
612	58	Tyr	CB	C	39.865	0.117	1
613	58	Tyr	CD1	C	133.148	0.030	1
614	58	Tyr	CD2	C	133.148	0.030	1
615	58	Tyr	CE1	C	118.053	0.036	1
616	58	Tyr	CE2	C	118.053	0.036	1
617	58	Tyr	H	H	8.317	0.005	1
618	58	Tyr	HA	H	5.221	0.004	1
619	58	Tyr	HB2	H	2.943	0.008	2
620	58	Tyr	HB3	H	2.697	0.015	2
621	58	Tyr	HD1	H	7.178	0.008	1
622	58	Tyr	HD2	H	7.178	0.008	1
623	58	Tyr	HE1	H	6.730	0.015	1
624	58	Tyr	HE2	H	6.730	0.015	1
625	58	Tyr	N	N	119.409	0.127	1
626	59	Val	C	C	174.760	0.000	1
627	59	Val	CA	C	59.764	0.062	1
628	59	Val	CB	C	35.727	0.085	1
629	59	Val	CG1	C	18.325	0.028	1
630	59	Val	CG2	C	22.809	0.072	1
631	59	Val	H	H	7.448	0.005	1
632	59	Val	HA	H	4.446	0.008	1
633	59	Val	HB	H	2.127	0.012	1
634	59	Val	HG11	H	0.491	0.011	1
635	59	Val	HG12	H	0.491	0.011	1
636	59	Val	HG13	H	0.491	0.011	1
637	59	Val	HG21	H	0.758	0.007	1
638	59	Val	HG22	H	0.758	0.007	1
639	59	Val	HG23	H	0.758	0.007	1
640	59	Val	N	N	115.047	0.106	1
641	60	Tyr	C	C	176.946	0.000	1
642	60	Tyr	CA	C	59.950	0.053	1
643	60	Tyr	CB	C	38.727	0.141	1
644	60	Tyr	CD1	C	133.289	0.014	1
645	60	Tyr	CD2	C	133.289	0.014	1
646	60	Tyr	CE1	C	118.351	0.018	1
647	60	Tyr	CE2	C	118.351	0.018	1
648	60	Tyr	H	H	8.326	0.005	1
649	60	Tyr	HA	H	4.531	0.006	1
650	60	Tyr	HB2	H	3.108	0.005	2
651	60	Tyr	HB3	H	2.962	0.006	2
652	60	Tyr	HD1	H	7.300	0.009	1
653	60	Tyr	HD2	H	7.300	0.009	1
654	60	Tyr	HE1	H	6.948	0.009	1
655	60	Tyr	HE2	H	6.948	0.009	1
656	60	Tyr	N	N	119.304	0.106	1
657	61	Gly	C	C	172.151	0.000	1
658	61	Gly	CA	C	43.631	0.072	1
659	61	Gly	H	H	8.152	0.005	1
660	61	Gly	HA2	H	4.017	0.007	2
661	61	Gly	HA3	H	3.805	0.008	2

662	61	Gly	N	N	109.696	0.091	1
663	62	Tyr	CA	C	57.220	0.064	1
664	62	Tyr	CB	C	38.625	0.031	1
665	62	Tyr	CD1	C	132.881	0.000	1
666	62	Tyr	CD2	C	132.881	0.000	1
667	62	Tyr	CE1	C	118.603	0.050	1
668	62	Tyr	CE2	C	118.603	0.050	1
669	62	Tyr	H	H	8.394	0.007	1
670	62	Tyr	HA	H	2.554	0.009	1
671	62	Tyr	HB2	H	2.621	0.033	2
672	62	Tyr	HB3	H	2.804	0.011	2
673	62	Tyr	HD1	H	7.018	0.029	1
674	62	Tyr	HD2	H	7.018	0.029	1
675	62	Tyr	HE1	H	6.904	0.006	1
676	62	Tyr	HE2	H	6.904	0.006	1
677	62	Tyr	N	N	123.921	0.087	1
678	63	Pro	C	C	174.657	0.000	1
679	63	Pro	CA	C	63.988	0.077	1
680	63	Pro	CB	C	32.549	0.087	1
681	63	Pro	CD	C	49.332	0.045	1
682	63	Pro	CG	C	23.897	0.063	1
683	63	Pro	HA	H	2.748	0.008	1
684	63	Pro	HB2	H	1.919	0.007	2
685	63	Pro	HB3	H	0.883	0.008	2
686	63	Pro	HD2	H	3.102	0.018	2
687	63	Pro	HD3	H	3.120	0.016	2
688	63	Pro	HG2	H	0.872	0.008	2
689	63	Pro	HG3	H	1.423	0.006	2
690	64	Tyr	C	C	173.768	0.000	1
691	64	Tyr	CA	C	60.681	0.050	1
692	64	Tyr	CB	C	35.178	0.047	1
693	64	Tyr	CD1	C	133.355	0.028	1
694	64	Tyr	CD2	C	133.355	0.028	1
695	64	Tyr	CE1	C	118.795	0.057	1
696	64	Tyr	CE2	C	118.795	0.057	1
697	64	Tyr	H	H	8.828	0.006	1
698	64	Tyr	HA	H	4.259	0.006	1
699	64	Tyr	HB2	H	3.449	0.013	2
700	64	Tyr	HB3	H	3.290	0.006	2
701	64	Tyr	HD1	H	6.983	0.021	1
702	64	Tyr	HD2	H	6.983	0.021	1
703	64	Tyr	HE1	H	6.886	0.007	1
704	64	Tyr	HE2	H	6.886	0.007	1
705	64	Tyr	N	N	118.836	0.076	1
706	65	Tyr	C	C	173.241	0.000	1
707	65	Tyr	CA	C	54.935	0.049	1
708	65	Tyr	CB	C	42.282	0.124	1
709	65	Tyr	CD1	C	133.835	0.030	1
710	65	Tyr	CD2	C	133.835	0.030	1
711	65	Tyr	CE1	C	117.818	0.000	1
712	65	Tyr	CE2	C	117.818	0.000	1
713	65	Tyr	H	H	6.912	0.010	1
714	65	Tyr	HA	H	5.716	0.006	1

715	65	Tyr	HB2	H	3.190	0.009	2
716	65	Tyr	HB3	H	2.617	0.014	2
717	65	Tyr	HD1	H	6.817	0.008	1
718	65	Tyr	HD2	H	6.817	0.008	1
719	65	Tyr	HE1	H	6.671	0.012	1
720	65	Tyr	HE2	H	6.671	0.012	1
721	65	Tyr	N	N	114.655	0.087	1
722	66	Lys	C	C	175.823	0.000	1
723	66	Lys	CA	C	55.020	0.050	1
724	66	Lys	CB	C	36.304	0.128	1
725	66	Lys	CE	C	42.231	0.019	1
726	66	Lys	CG	C	24.308	0.077	1
727	66	Lys	H	H	8.972	0.002	1
728	66	Lys	HA	H	4.441	0.007	1
729	66	Lys	HB2	H	1.597	0.017	2
730	66	Lys	HB3	H	1.570	0.018	2
731	66	Lys	HE2	H	2.903	0.004	1
732	66	Lys	HE3	H	2.903	0.004	1
733	66	Lys	HG2	H	1.253	0.011	2
734	66	Lys	HG3	H	1.144	0.019	2
735	66	Lys	N	N	119.044	0.106	1
736	67	Cys	C	C	174.677	0.000	1
737	67	Cys	CA	C	54.278	0.122	1
738	67	Cys	CB	C	42.401	0.176	1
739	67	Cys	H	H	9.334	0.002	1
740	67	Cys	HA	H	5.451	0.009	1
741	67	Cys	HB2	H	2.736	0.013	2
742	67	Cys	HB3	H	2.877	0.012	2
743	67	Cys	N	N	123.374	0.079	1
744	68	Ser	C	C	174.679	0.000	1
745	68	Ser	CA	C	57.363	0.028	1
746	68	Ser	CB	C	62.956	0.045	1
747	68	Ser	H	H	9.256	0.005	1
748	68	Ser	HA	H	4.775	0.013	1
749	68	Ser	HB2	H	3.814	0.020	2
750	68	Ser	HB3	H	3.922	0.007	2
751	68	Ser	N	N	121.159	0.111	1
752	69	Cys	CA	C	51.871	0.032	1
753	69	Cys	CB	C	37.256	0.063	1
754	69	Cys	H	H	9.090	0.004	1
755	69	Cys	HA	H	5.210	0.013	1
756	69	Cys	HB2	H	3.000	0.006	2
757	69	Cys	HB3	H	2.700	0.017	2
758	69	Cys	N	N	125.187	0.204	1
759	70	Pro	C	C	174.889	0.000	1
760	70	Pro	CA	C	62.302	0.069	1
761	70	Pro	CB	C	32.313	0.078	1
762	70	Pro	CD	C	50.700	0.029	1
763	70	Pro	CG	C	27.374	0.040	1
764	70	Pro	HA	H	4.705	0.006	1
765	70	Pro	HB2	H	2.454	0.005	2
766	70	Pro	HB3	H	2.142	0.009	2
767	70	Pro	HD2	H	3.262	0.006	2

768	70	Pro	HD3	H	3.522	0.004	2
769	70	Pro	HG2	H	2.154	0.004	2
770	70	Pro	HG3	H	2.008	0.005	2
771	71	Tyr	C	C	176.813	0.000	1
772	71	Tyr	CA	C	60.587	0.067	1
773	71	Tyr	CB	C	38.355	0.068	1
774	71	Tyr	CD1	C	133.327	0.035	1
775	71	Tyr	CD2	C	133.327	0.035	1
776	71	Tyr	CE1	C	118.205	0.032	1
777	71	Tyr	CE2	C	118.205	0.032	1
778	71	Tyr	H	H	8.452	0.005	1
779	71	Tyr	HA	H	4.235	0.008	1
780	71	Tyr	HB2	H	3.184	0.007	2
781	71	Tyr	HB3	H	2.884	0.007	2
782	71	Tyr	HD1	H	7.165	0.003	1
783	71	Tyr	HD2	H	7.165	0.003	1
784	71	Tyr	HE1	H	6.850	0.004	1
785	71	Tyr	HE2	H	6.850	0.004	1
786	71	Tyr	N	N	119.393	0.121	1
787	72	Gly	C	C	173.025	0.000	1
788	72	Gly	CA	C	44.825	0.052	1
789	72	Gly	H	H	8.290	0.010	1
790	72	Gly	HA2	H	3.942	0.009	2
791	72	Gly	HA3	H	3.065	0.008	2
792	72	Gly	N	N	113.221	0.076	1
793	73	Tyr	C	C	174.532	0.000	1
794	73	Tyr	CA	C	57.292	0.049	1
795	73	Tyr	CB	C	41.740	0.121	1
796	73	Tyr	CD1	C	133.787	0.041	1
797	73	Tyr	CD2	C	133.787	0.041	1
798	73	Tyr	CE1	C	117.759	0.039	1
799	73	Tyr	CE2	C	117.759	0.039	1
800	73	Tyr	H	H	7.993	0.003	1
801	73	Tyr	HA	H	5.561	0.010	1
802	73	Tyr	HB2	H	3.010	0.009	2
803	73	Tyr	HB3	H	2.646	0.007	2
804	73	Tyr	HD1	H	6.898	0.007	1
805	73	Tyr	HD2	H	6.898	0.007	1
806	73	Tyr	HE1	H	6.697	0.007	1
807	73	Tyr	HE2	H	6.697	0.007	1
808	73	Tyr	N	N	118.830	0.085	1
809	74	Tyr	C	C	173.299	0.000	1
810	74	Tyr	CA	C	56.838	0.082	1
811	74	Tyr	CB	C	38.418	0.081	1
812	74	Tyr	CD1	C	133.847	0.020	1
813	74	Tyr	CD2	C	133.847	0.020	1
814	74	Tyr	CE1	C	118.369	0.022	1
815	74	Tyr	CE2	C	118.369	0.022	1
816	74	Tyr	H	H	9.445	0.003	1
817	74	Tyr	HA	H	4.928	0.008	1
818	74	Tyr	HB2	H	3.083	0.012	2
819	74	Tyr	HB3	H	3.007	0.012	2
820	74	Tyr	HD1	H	6.869	0.006	1

821	74	Tyr	HD2	H	6.869	0.006	1
822	74	Tyr	HE1	H	6.497	0.005	1
823	74	Tyr	HE2	H	6.497	0.005	1
824	74	Tyr	N	N	119.356	0.087	1
825	75	Gly	C	C	176.797	0.000	1
826	75	Gly	CA	C	43.277	0.066	1
827	75	Gly	H	H	8.151	0.011	1
828	75	Gly	HA2	H	4.602	0.008	2
829	75	Gly	HA3	H	3.838	0.008	2
830	75	Gly	N	N	106.932	0.096	1
831	76	Lys	C	C	177.507	0.000	1
832	76	Lys	CA	C	59.708	0.088	1
833	76	Lys	CB	C	32.459	0.204	1
834	76	Lys	CD	C	29.342	0.065	1
835	76	Lys	CE	C	42.205	0.021	1
836	76	Lys	CG	C	25.693	0.030	1
837	76	Lys	H	H	9.277	0.008	1
838	76	Lys	HA	H	4.048	0.007	1
839	76	Lys	HB2	H	1.906	0.008	1
840	76	Lys	HB3	H	1.906	0.008	1
841	76	Lys	HD2	H	1.737	0.007	1
842	76	Lys	HD3	H	1.737	0.007	1
843	76	Lys	HE2	H	2.988	0.006	1
844	76	Lys	HE3	H	2.988	0.006	1
845	76	Lys	HG2	H	1.526	0.004	2
846	76	Lys	HG3	H	1.645	0.006	2
847	76	Lys	N	N	122.717	0.044	1
848	77	Gln	C	C	175.548	0.000	1
849	77	Gln	CA	C	55.551	0.063	1
850	77	Gln	CB	C	28.949	0.105	1
851	77	Gln	CG	C	35.747	0.032	1
852	77	Gln	H	H	8.347	0.005	1
853	77	Gln	HA	H	5.116	0.006	1
854	77	Gln	HB2	H	2.775	0.008	2
855	77	Gln	HB3	H	1.821	0.012	2
856	77	Gln	HG2	H	2.445	0.004	2
857	77	Gln	HG3	H	2.114	0.008	2
858	77	Gln	N	N	113.519	0.092	1
859	78	Cys	C	C	174.255	0.000	1
860	78	Cys	CA	C	54.839	0.047	1
861	78	Cys	CB	C	35.302	0.072	1
862	78	Cys	H	H	7.738	0.011	1
863	78	Cys	HA	H	4.035	0.009	1
864	78	Cys	HB2	H	3.699	0.013	2
865	78	Cys	HB3	H	2.963	0.005	2
866	78	Cys	N	N	111.537	0.076	1
867	79	Gln	C	C	175.337	0.000	1
868	79	Gln	CA	C	56.626	0.058	1
869	79	Gln	CB	C	28.252	0.123	1
870	79	Gln	CG	C	32.802	0.037	1
871	79	Gln	H	H	9.730	0.004	1
872	79	Gln	HA	H	4.075	0.009	1
873	79	Gln	HB2	H	2.192	0.009	2

874	79	Gln	HB3	H	1.880	0.006	2
875	79	Gln	HG2	H	2.755	0.004	2
876	79	Gln	HG3	H	2.225	0.004	2
877	79	Gln	N	N	118.080	0.104	1
878	80	Leu	C	C	174.419	0.000	1
879	80	Leu	CA	C	53.459	0.079	1
880	80	Leu	CB	C	41.891	0.066	1
881	80	Leu	CD1	C	25.428	0.023	2
882	80	Leu	CD2	C	22.523	0.073	2
883	80	Leu	CG	C	26.701	0.039	1
884	80	Leu	H	H	8.539	0.010	1
885	80	Leu	HA	H	4.566	0.007	1
886	80	Leu	HB2	H	0.410	0.025	2
887	80	Leu	HB3	H	0.736	0.011	2
888	80	Leu	HD11	H	0.590	0.012	2
889	80	Leu	HD12	H	0.590	0.012	2
890	80	Leu	HD13	H	0.590	0.012	2
891	80	Leu	HD21	H	0.800	0.014	2
892	80	Leu	HD22	H	0.800	0.014	2
893	80	Leu	HD23	H	0.800	0.014	2
894	80	Leu	HG	H	1.188	0.006	1
895	80	Leu	N	N	124.626	0.111	1
896	81	Lys	C	C	176.317	0.000	1
897	81	Lys	CA	C	55.118	0.041	1
898	81	Lys	CB	C	34.800	0.052	1
899	81	Lys	CD	C	29.365	0.039	1
900	81	Lys	CE	C	42.265	0.021	1
901	81	Lys	CG	C	25.071	0.120	1
902	81	Lys	H	H	8.303	0.004	1
903	81	Lys	HA	H	4.457	0.007	1
904	81	Lys	HB2	H	1.610	0.006	2
905	81	Lys	HB3	H	1.456	0.009	2
906	81	Lys	HD2	H	1.511	0.014	1
907	81	Lys	HD3	H	1.511	0.014	1
908	81	Lys	HE2	H	2.847	0.004	1
909	81	Lys	HE3	H	2.847	0.004	1
910	81	Lys	HG2	H	1.017	0.008	2
911	81	Lys	HG3	H	0.886	0.006	2
912	81	Lys	N	N	121.789	0.062	1
913	82	Lys	C	C	174.957	0.000	1
914	82	Lys	CA	C	56.749	0.049	1
915	82	Lys	CB	C	33.637	0.070	1
916	82	Lys	CD	C	29.056	0.075	1
917	82	Lys	CE	C	41.853	0.036	1
918	82	Lys	CG	C	25.131	0.061	1
919	82	Lys	H	H	8.323	0.004	1
920	82	Lys	HA	H	4.151	0.007	1
921	82	Lys	HB2	H	1.414	0.012	2
922	82	Lys	HB3	H	1.272	0.006	2
923	82	Lys	HD2	H	1.273	0.006	1
924	82	Lys	HD3	H	1.273	0.006	1
925	82	Lys	HE2	H	2.576	0.003	1
926	82	Lys	HE3	H	2.576	0.003	1

927	82	Lys	HG2	H	0.993	0.007	1
928	82	Lys	HG3	H	0.993	0.007	1
929	82	Lys	N	N	125.768	0.056	1
930	83	Tyr	CA	C	58.858	0.052	1
931	83	Tyr	CB	C	39.483	0.061	1
932	83	Tyr	CD1	C	133.281	0.058	1
933	83	Tyr	CD2	C	133.281	0.058	1
934	83	Tyr	CE1	C	117.999	0.000	1
935	83	Tyr	CE2	C	117.999	0.000	1
936	83	Tyr	H	H	7.694	0.006	1
937	83	Tyr	HA	H	4.367	0.007	1
938	83	Tyr	HB2	H	3.077	0.006	2
939	83	Tyr	HB3	H	2.801	0.008	2
940	83	Tyr	HD1	H	7.061	0.004	1
941	83	Tyr	HD2	H	7.061	0.004	1
942	83	Tyr	HE1	H	6.715	0.021	1
943	83	Tyr	HE2	H	6.715	0.021	1
944	83	Tyr	N	N	125.799	0.051	1

ANNEX III.4: T1,T2, and heteroNOE analysis

Dynamic Center Analysis of T1 value at 600 MHz

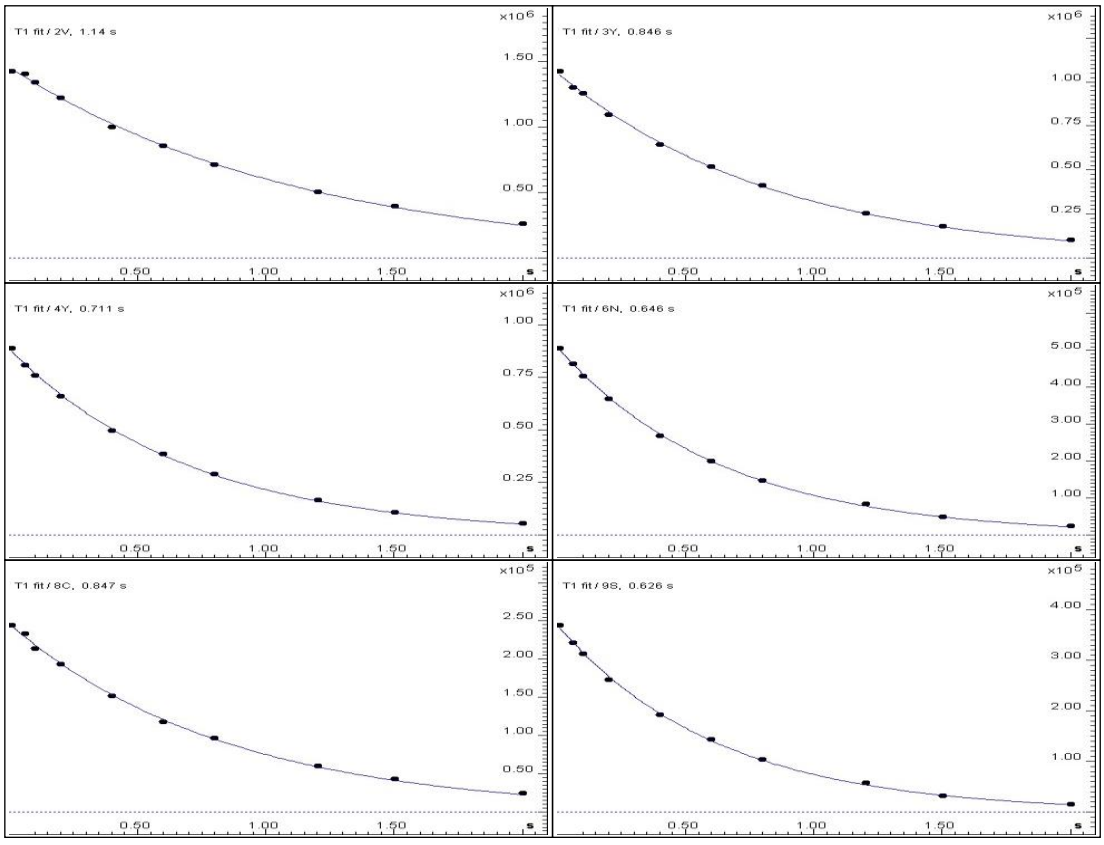
Fitted function:	f(t) = Io * exp (-t/T)
Random error estimation of data:	RMS per spectrum (or trace/plane)
Systematic error estimation of data:	worst case per peak scenario
Fit parameter Error estimation method:	from fit using arbitray y uncertainties
Confidence level:	95%
Used peaks:	peaks from D:\nmrdata\ToCompare\2020.11.10_15N13C_Pvfp5b_dynamics_600_298K\10007\pdata\1\peaklist.xml
Used integrals:	peak intensities
Used Mixing time:	all values (including replicates) used

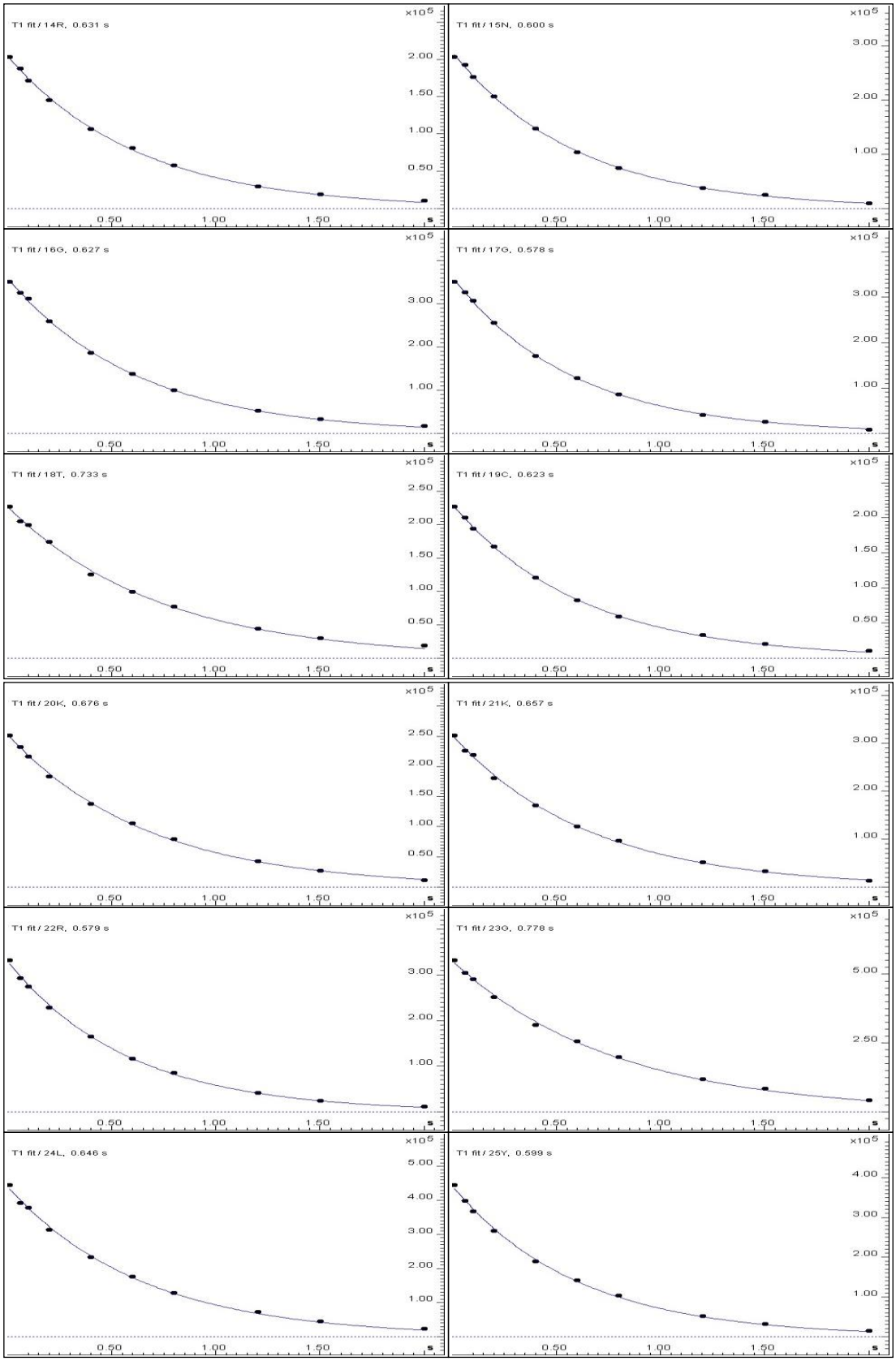
Peak name	F1 [ppm]	F2 [ppm]	Io	error	T1 [s]	error	fitInfo
2V	119.313	8.251	1.46e+06	2.072e+04	1.14	0.04328	Done
3Y	124.689	8.163	1.05e+06	1.632e+04	0.846	0.03298	Done
4Y	125.189	7.887	8.85e+05	1.218e+04	0.711	0.02415	Done
6N	118.313	8.327	5.09e+05	5938	0.646	0.01844	Done
8C	112.187	7.758	2.46e+05	3982	0.847	0.03448	Done
9S	116.062	7.365	3.70e+05	5814	0.626	0.02404	Done
14R	122.814	8.509	2.04e+05	3595	0.631	0.02712	Done
15N	113.937	9.078	2.89e+05	3699	0.600	0.01873	Done
16G	106.936	8.761	3.60e+05	4055	0.627	0.01729	Done
17G	105.686	7.353	3.44e+05	3251	0.578	0.01327	Done
18T	119.188	8.838	2.27e+05	4904	0.733	0.03914	Done
19C	127.690	8.421	2.19e+05	2077	0.623	0.01442	Done
20K	130.315	9.237	2.53e+05	3393	0.676	0.02227	Done
21K	123.939	8.750	3.17e+05	5722	0.657	0.02910	Done
22R	125.939	8.245	3.30e+05	6325	0.579	0.02698	Done
23G	111.437	8.139	5.49e+05	9274	0.778	0.03277	Done
24L	119.438	8.151	4.41e+05	9777	0.646	0.03513	Done
25Y	111.562	8.010	3.80e+05	7566	0.599	0.02907	Done

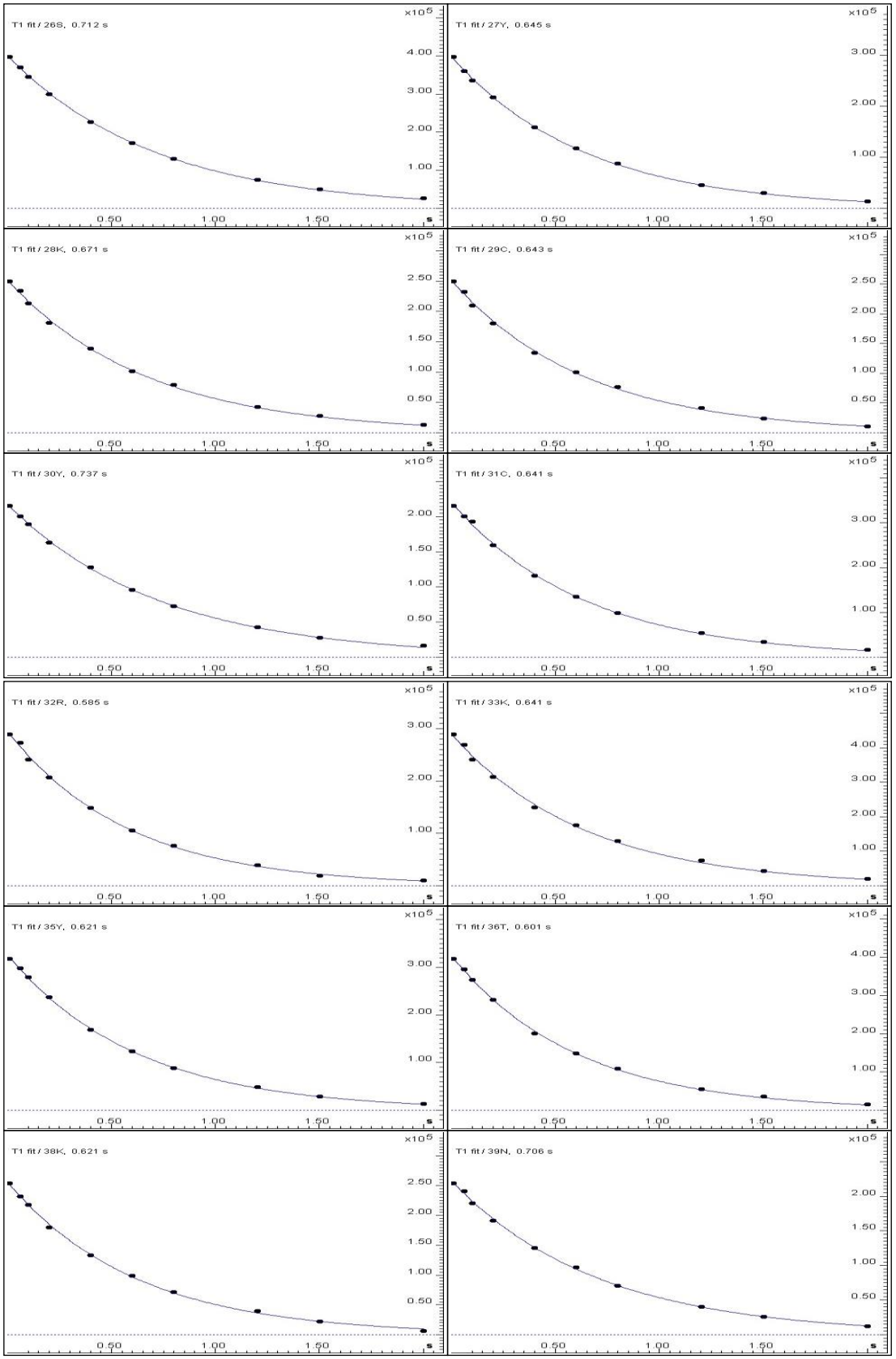
26S	113.687	7.177	4.01e+05	3165	0.712	0.01387	Done
27Y	117.438	7.928	2.97e+05	3675	0.645	0.01951	Done
28K	120.813	9.060	2.52e+05	4424	0.671	0.02892	Done
29C	121.688	9.043	2.56e+05	4853	0.643	0.02979	Done
30Y	124.939	9.354	2.18e+05	2007	0.737	0.01683	Done
31C	123.564	8.785	3.46e+05	4655	0.641	0.02108	Done
32R	120.938	8.415	2.95e+05	6094	0.585	0.02948	Done
33K	121.313	8.274	4.40e+05	1.025e+04	0.641	0.03653	Done
35Y	117.313	7.658	3.26e+05	2364	0.621	0.01099	Done
36T	112.562	8.914	4.04e+05	4162	0.601	0.01509	Done
38K	124.189	9.119	2.56e+05	4128	0.621	0.02446	Done
39N	111.687	8.362	2.22e+05	2887	0.706	0.02261	Done
40C	113.562	7.916	2.62e+05	2847	0.534	0.01405	Done
41Q	117.313	9.630	2.73e+05	1336	0.598	0.007134	Done
42Y	123.564	8.644	3.50e+05	4408	0.629	0.01938	Done
43N	121.188	8.843	3.78e+05	4304	0.649	0.01807	Done

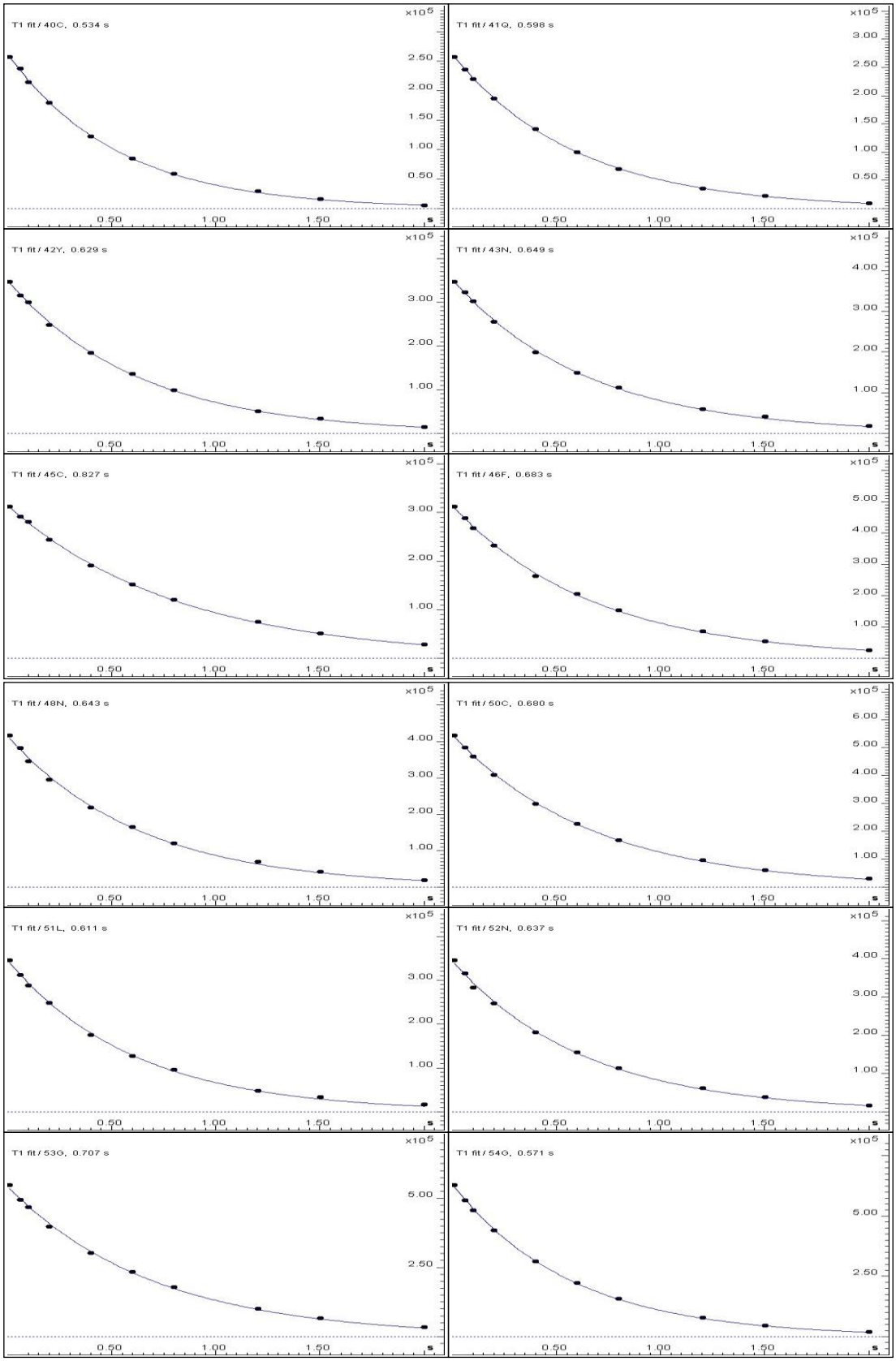
Peak name	F1 [ppm]	F2 [ppm]	Io	error	T1 [s]	error	fitInfo
45C	107.561	7.963	3.15e+05	2525	0.827	0.01663	Done
46F	121.313	7.335	4.87e+05	6339	0.683	0.02183	Done
48N	122.189	8.503	4.15e+05	9285	0.643	0.03516	Done
50C	118.688	7.670	5.48e+05	5361	0.680	0.01636	Done
51L	125.314	8.392	3.46e+05	6194	0.611	0.02672	Done
52N	110.687	8.509	3.95e+05	8636	0.637	0.03412	Done
53G	102.435	8.169	5.43e+05	1.038e+04	0.707	0.03333	Done
54G	106.811	7.658	6.32e+05	6716	0.571	0.01475	Done
55T	115.687	8.515	4.00e+05	2873	0.698	0.01237	Done
56C	128.065	9.008	3.76e+05	2778	0.614	0.01106	Done
57G	118.063	8.779	4.09e+05	8708	0.596	0.03087	Done
58	112.562	8.315	3.28e+05	4380	0.608	0.01981	Done
59V	114.687	7.353	3.60e+05	6081	0.607	0.02502	Done
61G	109.436	8.063	7.81e+05	1.201e+04	0.587	0.02200	Done
62Y	123.689	8.304	6.23e+05	6836	0.666	0.01794	Done
64Y	118.563	8.732	3.96e+05	9135	0.555	0.03106	Done

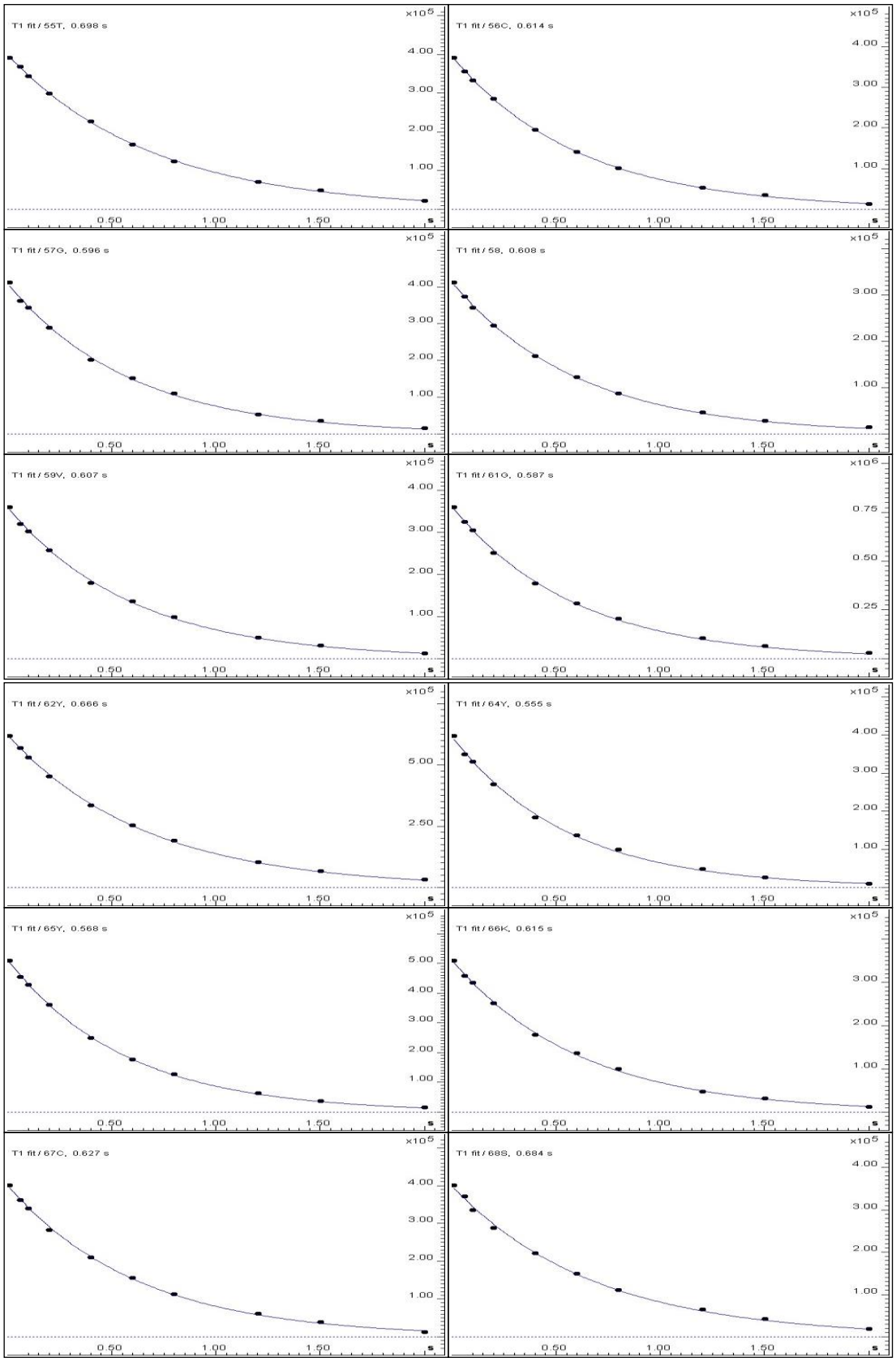
65Y	114.312	6.813	5.12e+05	5983	0.568	0.01612	Done
66K	118.813	8.873	3.52e+05	5546	0.615	0.02366	Done
67C	123.189	9.242	4.01e+05	6625	0.627	0.02530	Done
68S	120.813	9.160	3.57e+05	8187	0.684	0.03860	Done
69C	124.939	8.996	5.60e+05	5411	0.651	0.01540	Done
71Y	119.188	8.356	6.05e+05	1.016e+04	0.654	0.02691	Done
72G	112.937	8.192	4.82e+05	8529	0.645	0.02797	Done
73Y	118.563	7.893	5.61e+05	3807	0.604	0.009998	Done
74Y	119.063	9.348	4.85e+05	7927	0.607	0.02417	Done
76K	122.439	9.184	3.99e+05	8928	0.608	0.03315	Done
77Q	113.312	8.245	3.99e+05	4899	0.668	0.02014	Done
78C	111.312	7.641	4.19e+05	6006	0.569	0.01983	Done
79Q	117.813	9.635	4.05e+05	3330	0.629	0.01262	Done
80L	124.314	8.444	4.27e+05	8202	0.606	0.02838	Done
81K	121.563	8.210	5.60e+05	7759	0.656	0.02228	Done
82K	125.564	8.227	5.82e+05	9567	0.643	0.02588	Done
83Y	125.564	7.599	9.63e+05	1.224e+04	0.725	0.02275	Done

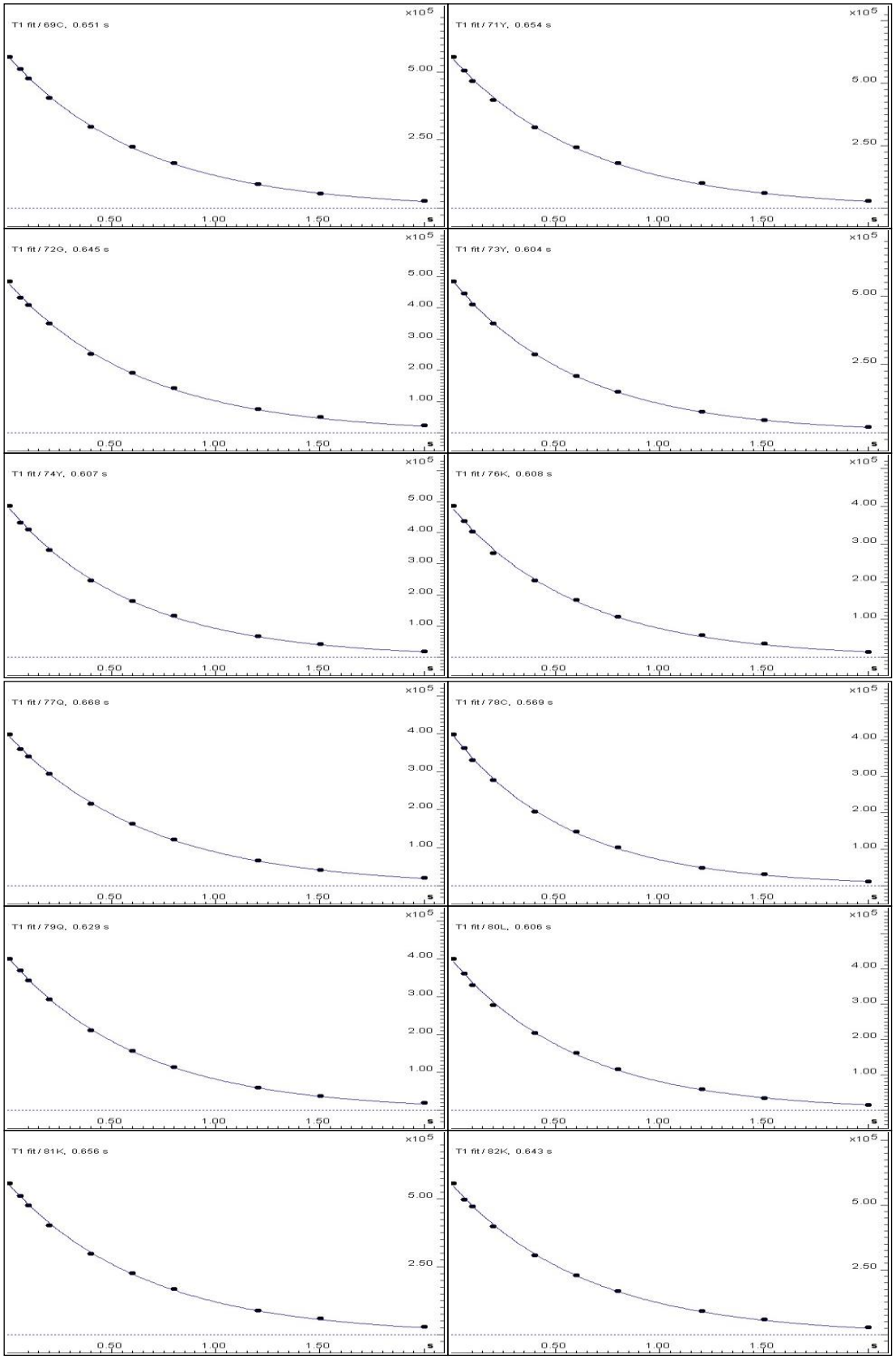


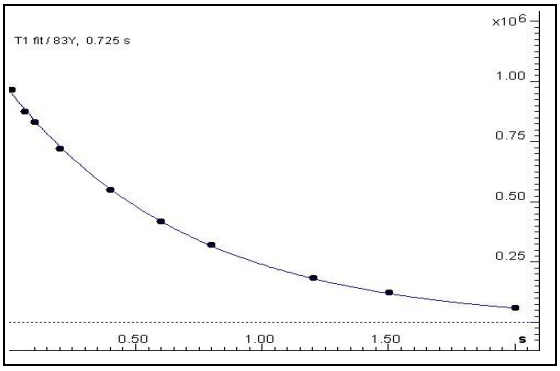

T1 fit / 14R, 0.631 s
T1 fit / 15N, 0.600 s
T1 fit / 16G, 0.627 s
T1 fit / 17G, 0.578 s
T1 fit / 18T, 0.733 s
T1 fit / 19C, 0.623 s
T1 fit / 20K, 0.676 s
T1 fit / 21K, 0.657 s
T1 fit / 22R, 0.579 s
T1 fit / 23G, 0.778 s
T1 fit / 24L, 0.646 s
T1 fit / 25Y, 0.599 s
x10^5
0.50
1.00
1.50
s


T1 fit / 26S, 0.712 s
T1 fit / 27Y, 0.645 s
T1 fit / 28K, 0.671 s
T1 fit / 29C, 0.643 s
T1 fit / 30Y, 0.737 s
T1 fit / 31C, 0.641 s
T1 fit / 32R, 0.585 s
T1 fit / 33K, 0.641 s
T1 fit / 35Y, 0.621 s
T1 fit / 36T, 0.601 s
T1 fit / 38K, 0.621 s
T1 fit / 39N, 0.706 s
x10^5
0.50
1.00
1.50
s


T1 fit / 40C, 0.534 s
T1 fit / 41Q, 0.598 s
T1 fit / 42Y, 0.629 s
T1 fit / 43N, 0.649 s
T1 fit / 45C, 0.827 s
T1 fit / 46F, 0.683 s
T1 fit / 48N, 0.643 s
T1 fit / 50C, 0.680 s
T1 fit / 51L, 0.611 s
T1 fit / 52N, 0.637 s
T1 fit / 53G, 0.707 s
T1 fit / 54G, 0.571 s
x10^5
0.50
1.00
1.50
s


T1 fit / 55T, 0.698 s
T1 fit / 56C, 0.614 s
T1 fit / 57G, 0.596 s
T1 fit / 58, 0.608 s
T1 fit / 59V, 0.607 s
T1 fit / 61G, 0.587 s
T1 fit / 62Y, 0.668 s
T1 fit / 64Y, 0.555 s
T1 fit / 65Y, 0.568 s
T1 fit / 66K, 0.615 s
T1 fit / 67C, 0.627 s
T1 fit / 68S, 0.684 s
s


T1 fit / 69C, 0.651 s
T1 fit / 71Y, 0.654 s
T1 fit / 72G, 0.645 s
T1 fit / 73Y, 0.604 s
T1 fit / 74Y, 0.607 s
T1 fit / 76K, 0.608 s
T1 fit / 77Q, 0.668 s
T1 fit / 78C, 0.569 s
T1 fit / 79Q, 0.629 s
T1 fit / 80L, 0.606 s
T1 fit / 81K, 0.656 s
T1 fit / 82K, 0.643 s
x10^5
0.50
1.00
1.50
s



Dynamic Center Analysis of T2 at 600 MHz

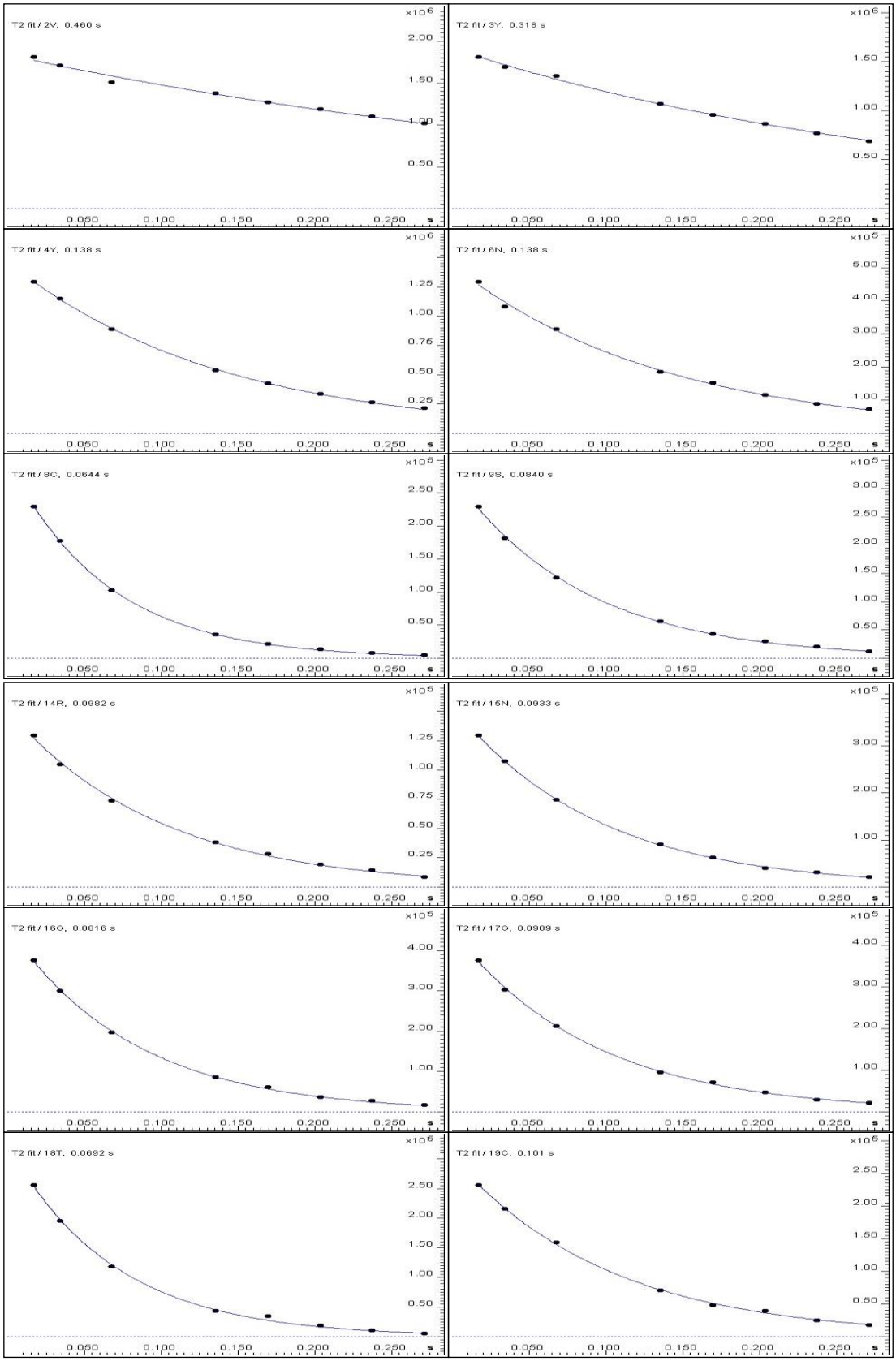
Fitted function:	f(t) = Io * exp (-t/T)
Random error estimation of data:	RMS per spectrum (or trace/plane)
Systematic error estimation of data:	worst case per peak scenario
Fit parameter Error estimation method:	from fit using arbitray y uncertainties
Confidence level:	95%
Used peaks:	peaks from D:\nmrdata\ToCompare\2020.11.10_15N13C_Pvfp5b_dynamics_600_298K\10007\pdata\1\peaklist.xml
Used integrals:	peak intensities
Used Time:	all values (including replicates) used

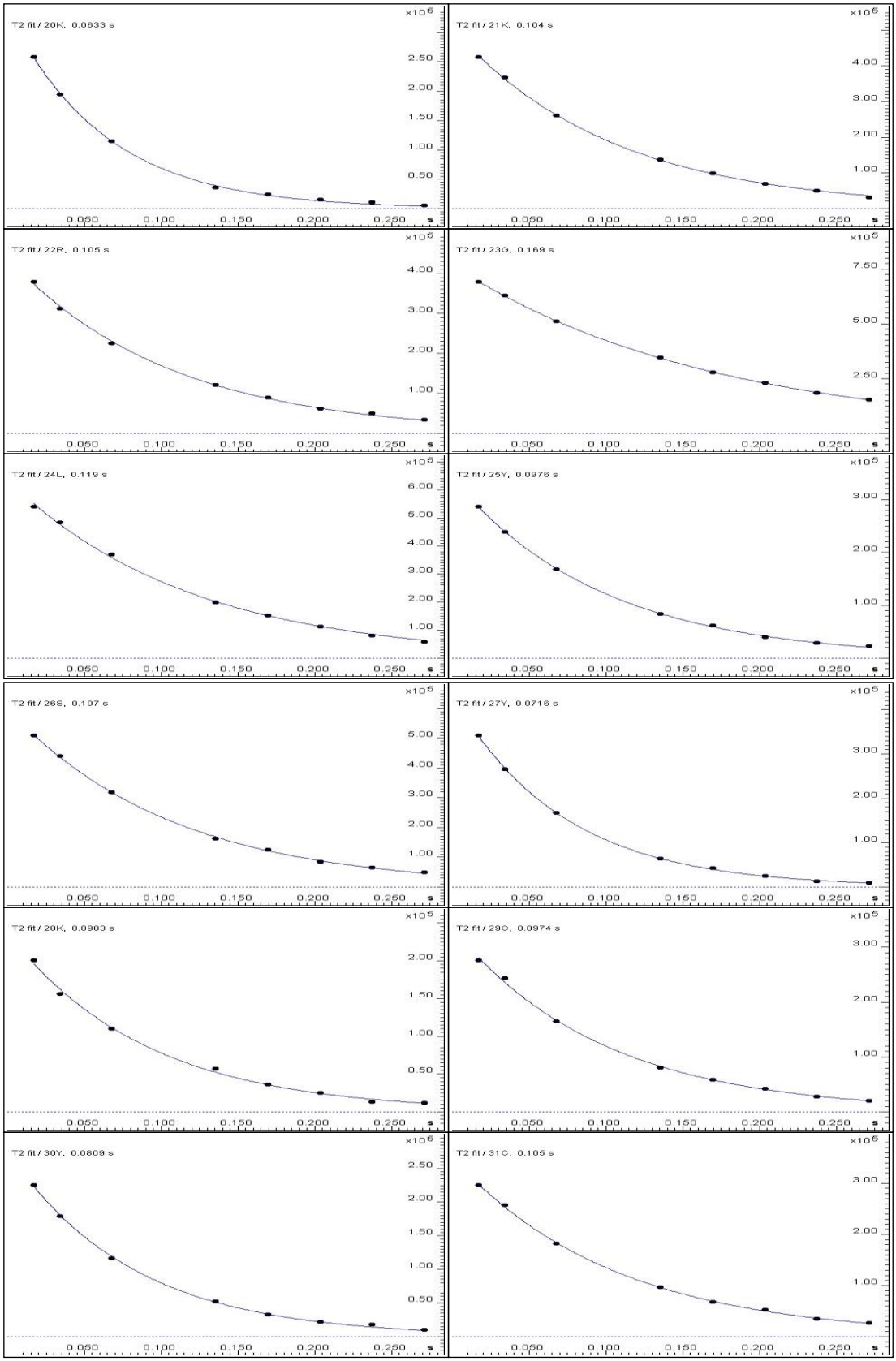
Peak name	F1 [ppm]	F2 [ppm]	T2 [s]	error	fitInfo
2V	119.376	8.257	0.460	0.05549	Done
3Y	124.627	8.163	0.318	0.01637	Done
4Y	125.127	7.887	0.138	0.002758	Done
6N	118.375	8.327	0.138	0.008411	Done
8C	112.124	7.758	0.0644	0.001363	Done
9S	115.875	7.365	0.0840	0.003432	Done
14R	122.626	8.509	0.0982	0.005318	Done
15N	113.875	9.072	0.0933	0.001744	Done
16G	106.873	8.761	0.0816	0.002741	Done
17G	105.623	7.353	0.0909	0.003652	Done
18T	119.126	8.838	0.0692	0.004589	Done
19C	127.627	8.415	0.101	0.004258	Done
20K	130.378	9.237	0.0633	0.002785	Done

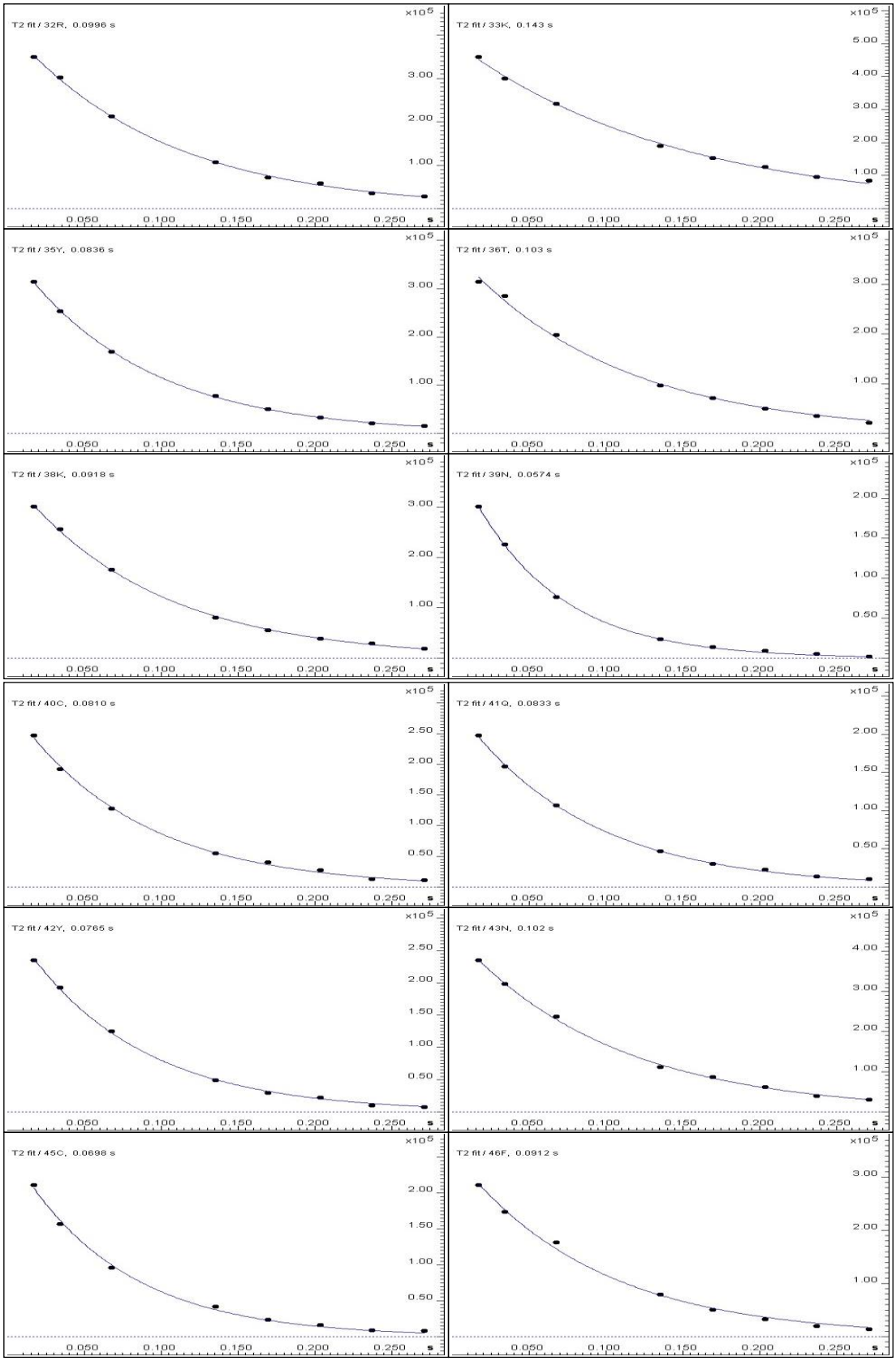
21K	123.876	8.750	0.104	0.002770	Done
22R	125.877	8.245	0.105	0.004719	Done
23G	111.374	8.139	0.169	0.001412	Done
24L	119.376	8.151	0.119	0.005792	Done
25Y	111.374	8.004	0.0976	0.002396	Done
26S	113.625	7.177	0.107	0.002983	Done
27Y	117.375	7.928	0.0716	0.001970	Done
28K	120.876	9.060	0.0903	0.007637	Done
29C	121.626	9.043	0.0974	0.005083	Done
30Y	124.877	9.354	0.0809	0.003396	Done
31C	123.376	8.785	0.105	0.002681	Done
32R	120.876	8.409	0.0996	0.004108	Done
33K	121.376	8.274	0.143	0.007921	Done
35Y	117.375	7.664	0.0836	0.001721	Done
36T	112.624	8.914	0.103	0.008121	Done
38K	124.126	9.113	0.0918	0.002844	Done
39N	111.624	8.362	0.0574	0.002042	Done
40C	113.375	7.910	0.0810	0.005239	Done
41Q	117.375	9.630	0.0833	0.003104	Done
42Y	123.626	8.644	0.0765	0.003636	Done
43N	121.126	8.843	0.102	0.004641	Done

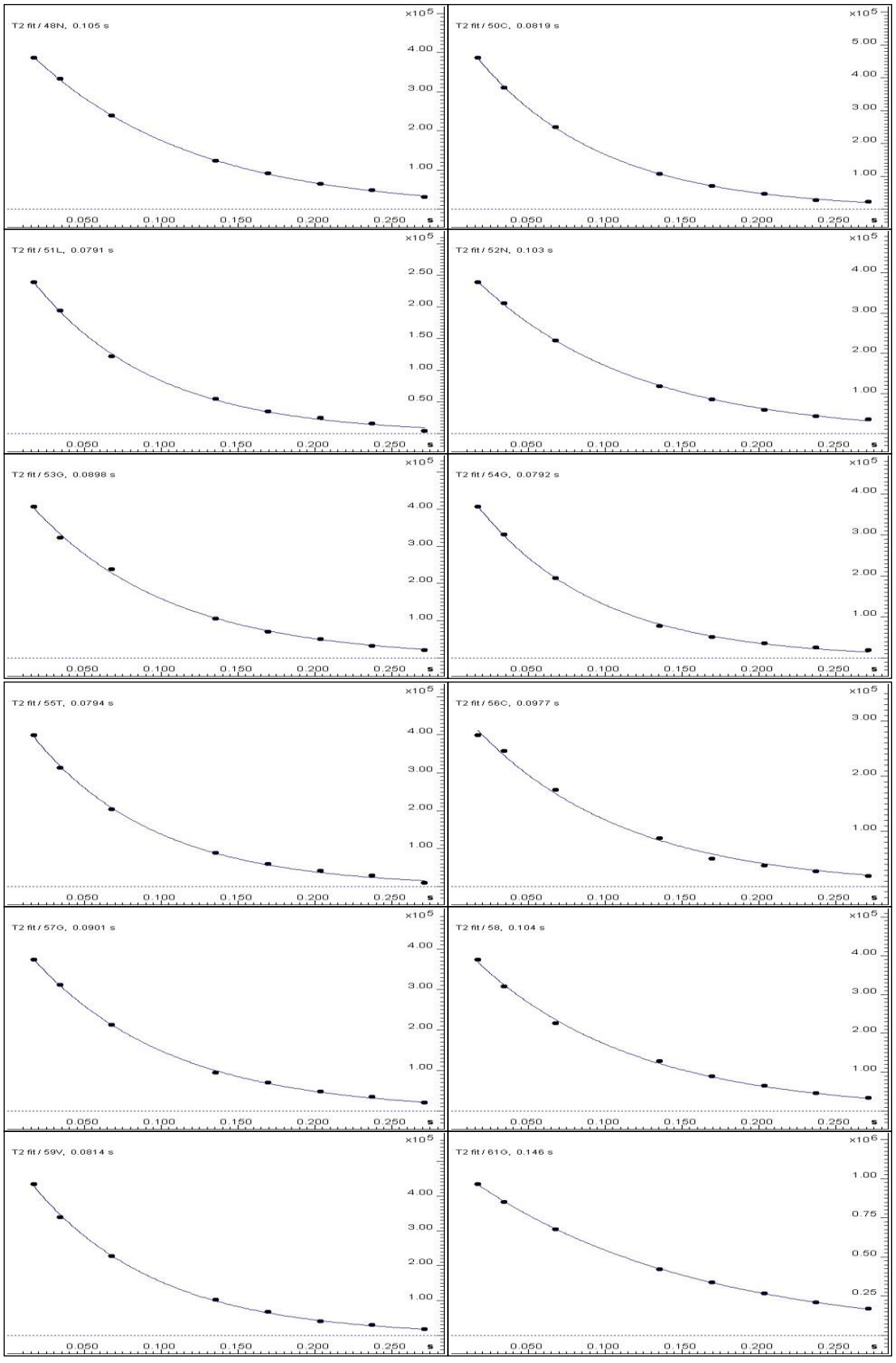
Peak name	F1 [ppm]	F2 [ppm]	T2 [s]	error	fitInfo
45C	107.373	7.963	0.0698	0.005940	Done
46F	121.376	7.335	0.0912	0.008257	Done
48N	122.126	8.503	0.105	0.002063	Done
50C	118.625	7.670	0.0819	0.002111	Done
51L	125.127	8.392	0.0791	0.004410	Done
52N	110.624	8.509	0.103	0.002349	Done
53G	102.373	8.169	0.0898	0.005593	Done
54G	106.623	7.658	0.0792	0.002676	Done
55T	115.625	8.509	0.0794	0.004227	Done
56C	127.877	9.008	0.0977	0.009343	Done
57G	117.875	8.773	0.0901	0.002570	Done

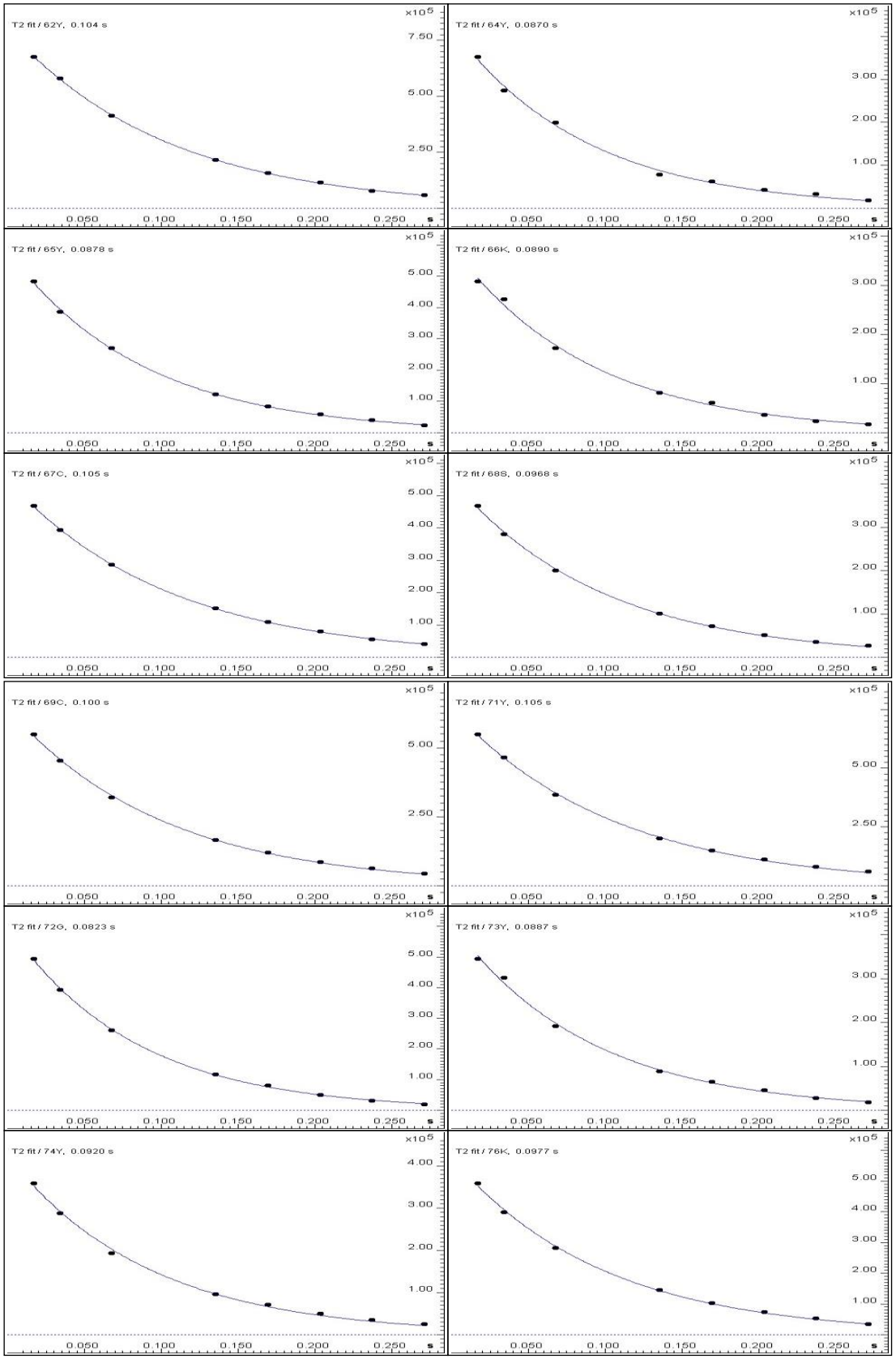
58	112.374	8.315	0.104	0.005968	Done
59V	114.625	7.347	0.0814	0.004057	Done
61G	109.374	8.063	0.146	0.001656	Done
62Y	123.626	8.304	0.104	0.001653	Done
64Y	118.625	8.732	0.0870	0.008040	Done
65Y	114.375	6.813	0.0878	0.003521	Done
66K	118.625	8.873	0.0890	0.007203	Done
67C	123.126	9.242	0.105	0.001909	Done
68S	120.876	9.160	0.0968	0.003704	Done
69C	124.877	8.996	0.100	0.003244	Done
71Y	119.126	8.356	0.105	0.003403	Done
72G	112.874	8.192	0.0823	0.002782	Done
73Y	118.375	7.893	0.0887	0.006517	Done
74Y	119.126	9.348	0.0920	0.006248	Done
76K	122.376	9.184	0.0977	0.004567	Done
77Q	113.124	8.245	0.0742	0.005075	Done
78C	111.124	7.641	0.0895	0.002564	Done
79Q	117.875	9.635	0.0799	0.004101	Done
80L	124.126	8.439	0.0847	0.005500	Done
81K	121.376	8.210	0.0999	0.007241	Done
82K	125.377	8.227	0.0865	0.001321	Done
83Y	125.377	7.599	0.114	0.003235	Done

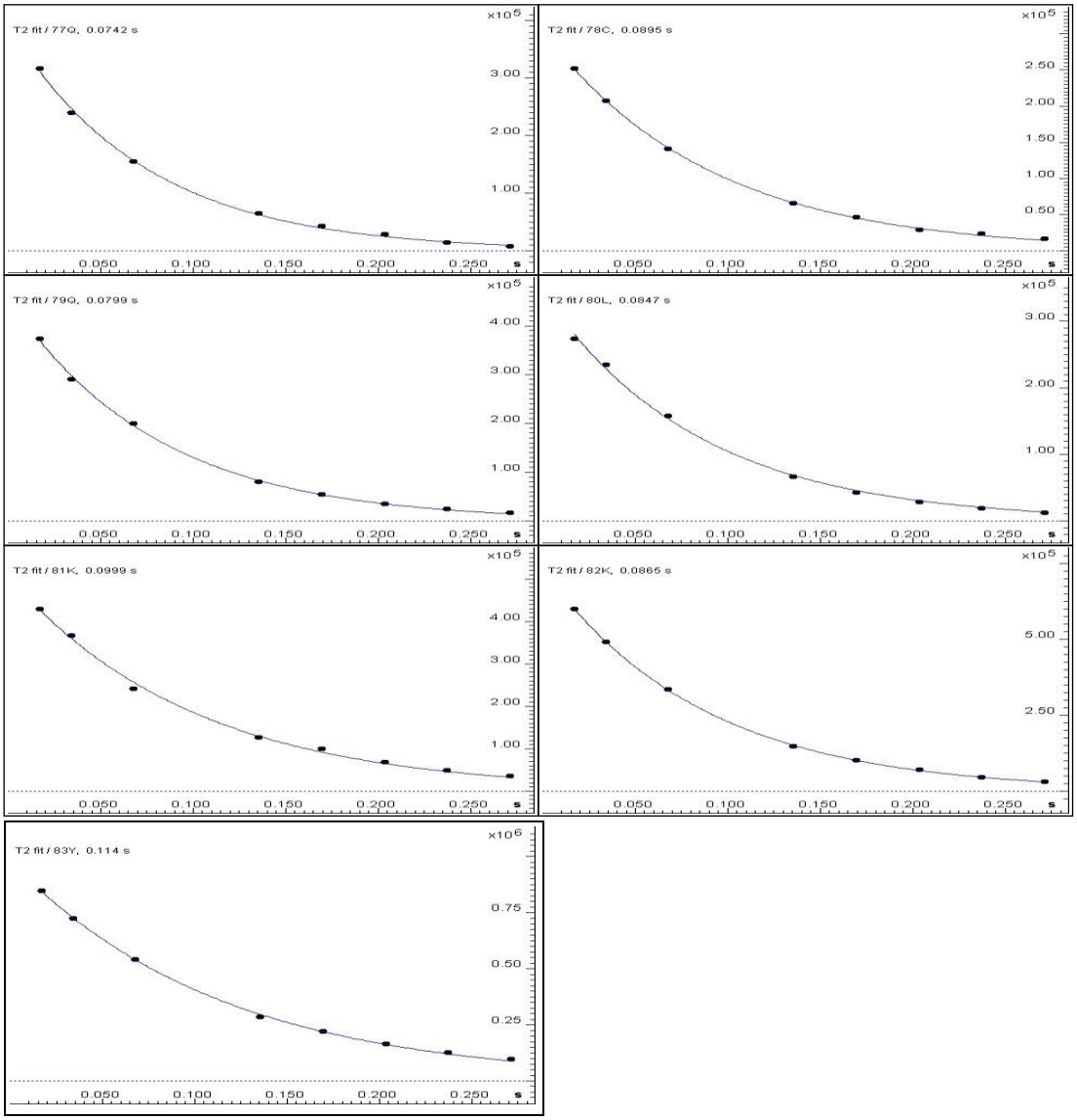

T2 fit / 2V, 0.460 s
T2 fit / 3Y, 0.318 s
T2 fit / 4Y, 0.138 s
T2 fit / 6N, 0.138 s
T2 fit / 8C, 0.0644 s
T2 fit / 9S, 0.0840 s
T2 fit / 14R, 0.0982 s
T2 fit / 15N, 0.0933 s
T2 fit / 16G, 0.0816 s
T2 fit / 17G, 0.0909 s
T2 fit / 18T, 0.0692 s
T2 fit / 19C, 0.101 s
s


T2 fit / 20K, 0.0633 s
T2 fit / 21K, 0.104 s
T2 fit / 22R, 0.105 s
T2 fit / 23G, 0.169 s
T2 fit / 24L, 0.119 s
T2 fit / 25Y, 0.0976 s
T2 fit / 26S, 0.107 s
T2 fit / 27Y, 0.0716 s
T2 fit / 28K, 0.0903 s
T2 fit / 29C, 0.0974 s
T2 fit / 30Y, 0.0809 s
T2 fit / 31C, 0.105 s


T2 fit / 32R, 0.0996 s
T2 fit / 33K, 0.143 s
T2 fit / 35Y, 0.0836 s
T2 fit / 36T, 0.103 s
T2 fit / 38K, 0.0918 s
T2 fit / 39N, 0.0574 s
T2 fit / 40C, 0.0810 s
T2 fit / 41Q, 0.0833 s
T2 fit / 42Y, 0.0765 s
T2 fit / 43N, 0.102 s
T2 fit / 45C, 0.0698 s
T2 fit / 46F, 0.0912 s
x10^5
0.050
0.100
0.150
0.200
0.250
s


T2 fit / 48N, 0.105 s
T2 fit / 50C, 0.0819 s
T2 fit / 51L, 0.0791 s
T2 fit / 52N, 0.103 s
T2 fit / 53G, 0.0898 s
T2 fit / 54G, 0.0792 s
T2 fit / 55T, 0.0794 s
T2 fit / 56C, 0.0977 s
T2 fit / 57G, 0.0901 s
T2 fit / 58, 0.104 s
T2 fit / 59V, 0.0814 s
T2 fit / 61G, 0.146 s


T2 fit / 62Y, 0.104 s
T2 fit / 64Y, 0.0870 s
T2 fit / 65Y, 0.0878 s
T2 fit / 66K, 0.0890 s
T2 fit / 67C, 0.105 s
T2 fit / 68S, 0.0968 s
T2 fit / 69C, 0.100 s
T2 fit / 71Y, 0.105 s
T2 fit / 72G, 0.0823 s
T2 fit / 73Y, 0.0887 s
T2 fit / 74Y, 0.0920 s
T2 fit / 76K, 0.0977 s
x10^5
0.050
0.100
0.150
0.200
0.250
s



Dynamic Center Analysis of NOE at 600 MHz

Random error estimation of data:	RMS per spectrum (or trace/plane)
Systematic error estimation of data:	none
Used peaks:	peaks from peaklist.xml at spectrum
Used integrals:	peak intensities

Peak name	F1 [ppm]	F2 [ppm]	NOE [ ]	error	fitInfo
2V	119.313	8.251	-0.628	0.002871	Fail
3Y	124.689	8.163	-0.468	0.003293	Fail
4Y	125.189	7.887	-0.174	0.002649	Fail

6N	118.313	8.327	0.399	0.005904	Fail
8C	112.187	7.758	0.676	0.01444	Fail
9S	116.062	7.365	0.731	0.009655	Fail
14R	122.814	8.503	0.703	0.01567	Fail
15N	113.937	9.072	0.770	0.01309	Fail
16G	106.936	8.761	0.795	0.01107	Fail
17G	105.686	7.353	0.771	0.01077	Fail
18T	119.188	8.838	0.707	0.01539	Fail
19C	127.690	8.415	0.738	0.01515	Fail
20K	130.315	9.231	0.725	0.01462	Fail
21K	123.939	8.750	0.690	0.01109	Fail
22R	125.939	8.245	0.658	0.01085	Fail
23G	111.437	8.139	0.228	0.004399	Fail
24L	119.438	8.151	0.584	0.008148	Fail
25Y	111.562	8.004	0.532	0.008290	Fail
26S	113.687	7.171	0.532	0.007824	Fail
27Y	117.438	7.922	0.713	0.01233	Fail
28K	120.813	9.060	0.723	0.01364	Fail
29C	121.688	9.043	0.730	0.01435	Fail
30Y	124.939	9.354	0.769	0.01663	Fail
31C	123.564	8.785	0.708	0.01059	Fail
32R	120.938	8.409	0.773	0.01202	Fail
33K	121.313	8.274	0.633	0.007854	Fail
35Y	117.313	7.658	0.756	0.01240	Fail
36T	112.562	8.914	0.737	0.01012	Fail
38K	124.189	9.113	0.772	0.01389	Fail
39N	111.687	8.362	0.744	0.01607	Fail
40C	113.562	7.910	0.783	0.01509	Fail
41Q	117.313	9.624	0.759	0.01312	Fail
42Y	123.564	8.644	0.778	0.01063	Fail
43N	121.188	8.843	0.764	0.01058	Fail
45C	107.561	7.957	0.734	0.01094	Fail
46F	121.313	7.330	0.739	0.007300	Fail
48N	122.189	8.497	0.733	0.008537	Fail

50C	118.688	7.670	0.742	0.007834	Fail
51L	125.314	8.392	0.757	0.01085	Fail

Peak name	F1 [ppm]	F2 [ppm]	NOE [ ]	error	fitInfo
52N	110.687	8.509	0.744	0.008853	Fail
53G	102.435	8.169	0.762	0.007541	Fail
54G	106.811	7.658	0.739	0.006608	Fail
55T	115.687	8.509	0.769	0.009211	Fail
56C	128.065	9.002	0.789	0.01104	Fail
57G	118.063	8.779	0.739	0.009240	Fail
58	112.437	8.315	0.647	0.01054	Fail
59V	114.687	7.347	0.736	0.009264	Fail
61G	109.436	8.063	0.532	0.004423	Fail
62Y	123.689	8.304	0.726	0.006742	Fail
64Y	118.563	8.726	0.709	0.008592	Fail
65Y	114.312	6.807	0.740	0.007343	Fail
66K	118.813	8.873	0.735	0.01130	Fail
67C	123.189	9.242	0.722	0.009097	Fail
68S	120.813	9.154	0.733	0.009835	Fail
69C	124.939	8.996	0.739	0.007622	Fail
71Y	119.188	8.356	0.667	0.005872	Fail
72G	112.937	8.192	0.651	0.007488	Fail
73Y	118.563	7.893	0.702	0.006747	Fail
74Y	119.063	9.348	0.743	0.007759	Fail
76K	122.439	9.178	0.697	0.009019	Fail
77Q	113.312	8.251	0.735	0.009388	Fail
78C	111.312	7.641	0.745	0.009486	Fail
79Q	117.813	9.630	0.746	0.009161	Fail
80L	124.314	8.439	0.759	0.008930	Fail
81K	121.563	8.210	0.676	0.006588	Fail
82K	125.564	8.227	0.570	0.005721	Fail
83Y	125.564	7.599	0.270	0.002542	Fail

Dynamic Center Analysis of T1 at 800 MHz

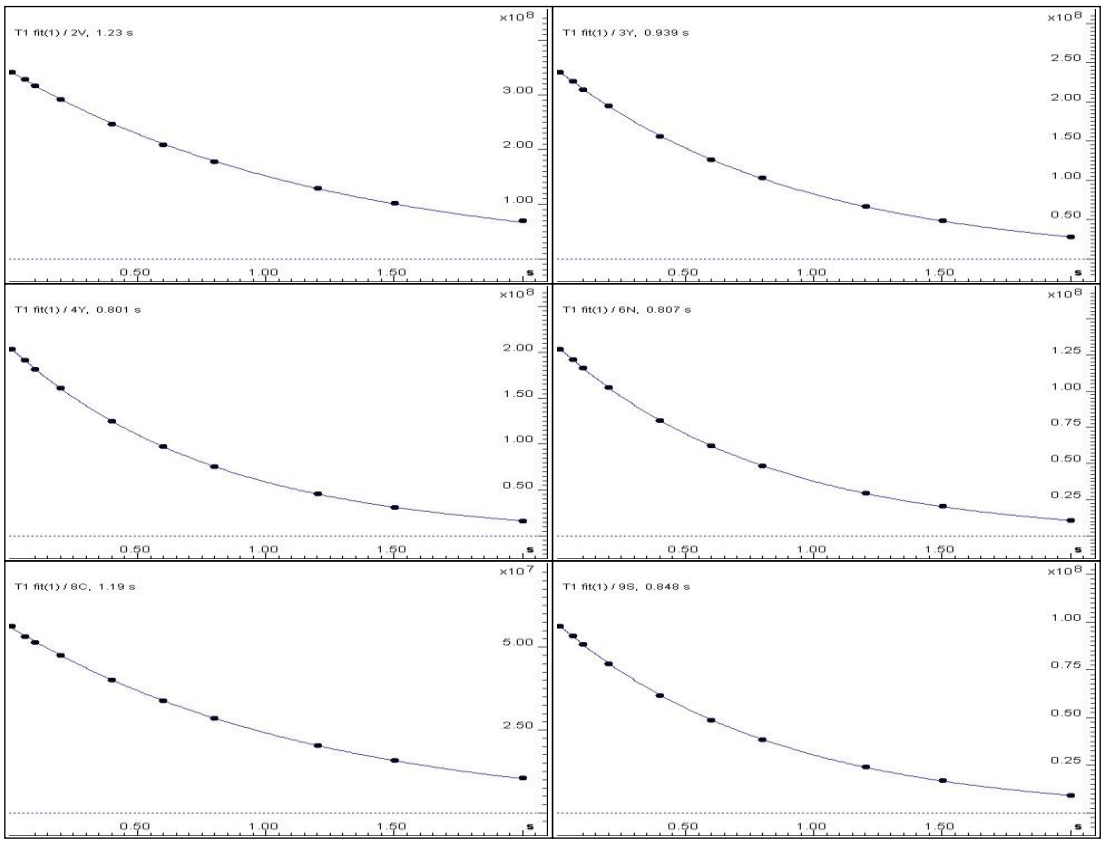
Fitted function:	f(t) = Io * exp (-t/T)
Random error estimation of data:	RMS per spectrum (or trace/plane)
Systematic error estimation of data:	worst case per peak scenario
Fit parameter Error estimation method:	from fit using arbitray y uncertainties
Confidence level:	95%
Used peaks:	peaks from peaklist.xml at spectrum
Used integrals:	peak intensities
Used Mixing time:	integrals related to replicate values replaced by mean

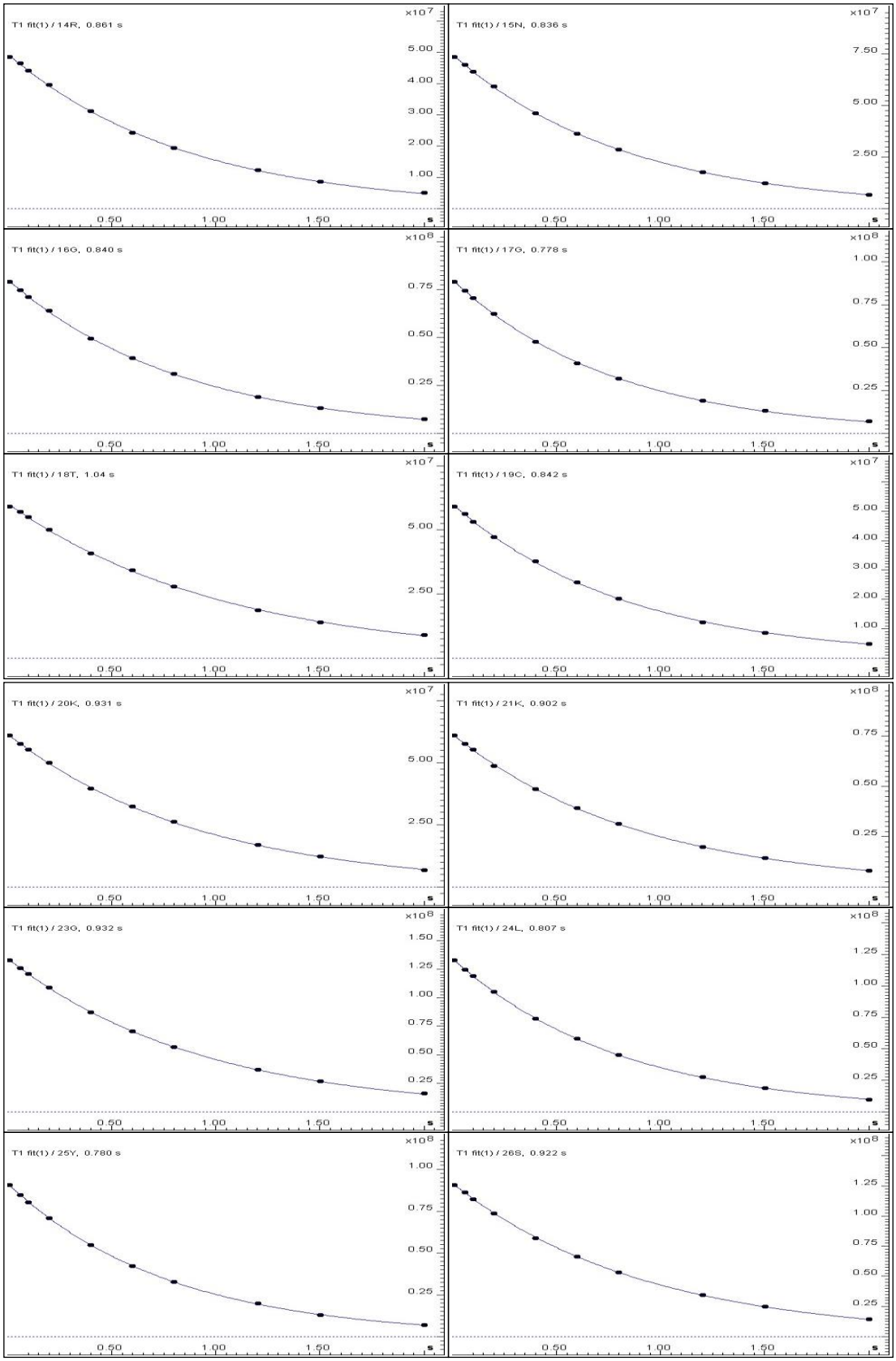
Peak name	F1 [ppm]	F2 [ppm]	Io	error	T1 [s]	error	fitInfo
2V	119.344	8.260	3.44e+08	2.123e+06	1.23	0.02069	Done
3Y	124.720	8.166	2.41e+08	7.101e+05	0.939	0.007086	Done
4Y	125.220	7.890	2.06e+08	4.387e+05	0.801	0.004256	Done
6N	118.344	8.330	1.31e+08	3.578e+05	0.807	0.005507	Done
8C	112.218	7.767	5.63e+07	3.726e+05	1.19	0.02122	Done
9S	116.094	7.374	9.94e+07	3.031e+05	0.848	0.006517	Done
14R	122.845	8.512	4.96e+07	3.518e+05	0.861	0.01541	Done
15N	113.968	9.076	7.49e+07	2.310e+05	0.836	0.006486	Done
16G	106.967	8.765	8.04e+07	4.403e+05	0.840	0.01157	Done
17G	105.655	7.362	8.98e+07	3.755e+05	0.778	0.008102	Done
18T	119.157	8.841	6.05e+07	5.053e+05	1.04	0.02282	Done
19C	127.783	8.424	5.25e+07	3.347e+05	0.842	0.01352	Done
20K	130.408	9.240	6.16e+07	3.365e+05	0.931	0.01301	Done
21K	123.970	8.753	7.61e+07	3.466e+05	0.902	0.01045	Done
23G	111.405	8.143	1.35e+08	3.354e+05	0.932	0.005943	Done
24L	119.532	8.154	1.22e+08	3.543e+05	0.807	0.005864	Done
25Y	111.593	8.014	9.17e+07	3.485e+05	0.780	0.007383	Done
26S	113.656	7.186	1.27e+08	4.446e+05	0.922	0.008208	Done
27Y	117.469	7.931	6.90e+07	4.194e+05	0.862	0.01322	Done
28K	120.844	9.064	6.17e+07	4.297e+05	0.939	0.01676	Done
29C	121.657	9.046	6.09e+07	3.405e+05	0.898	0.01278	Done

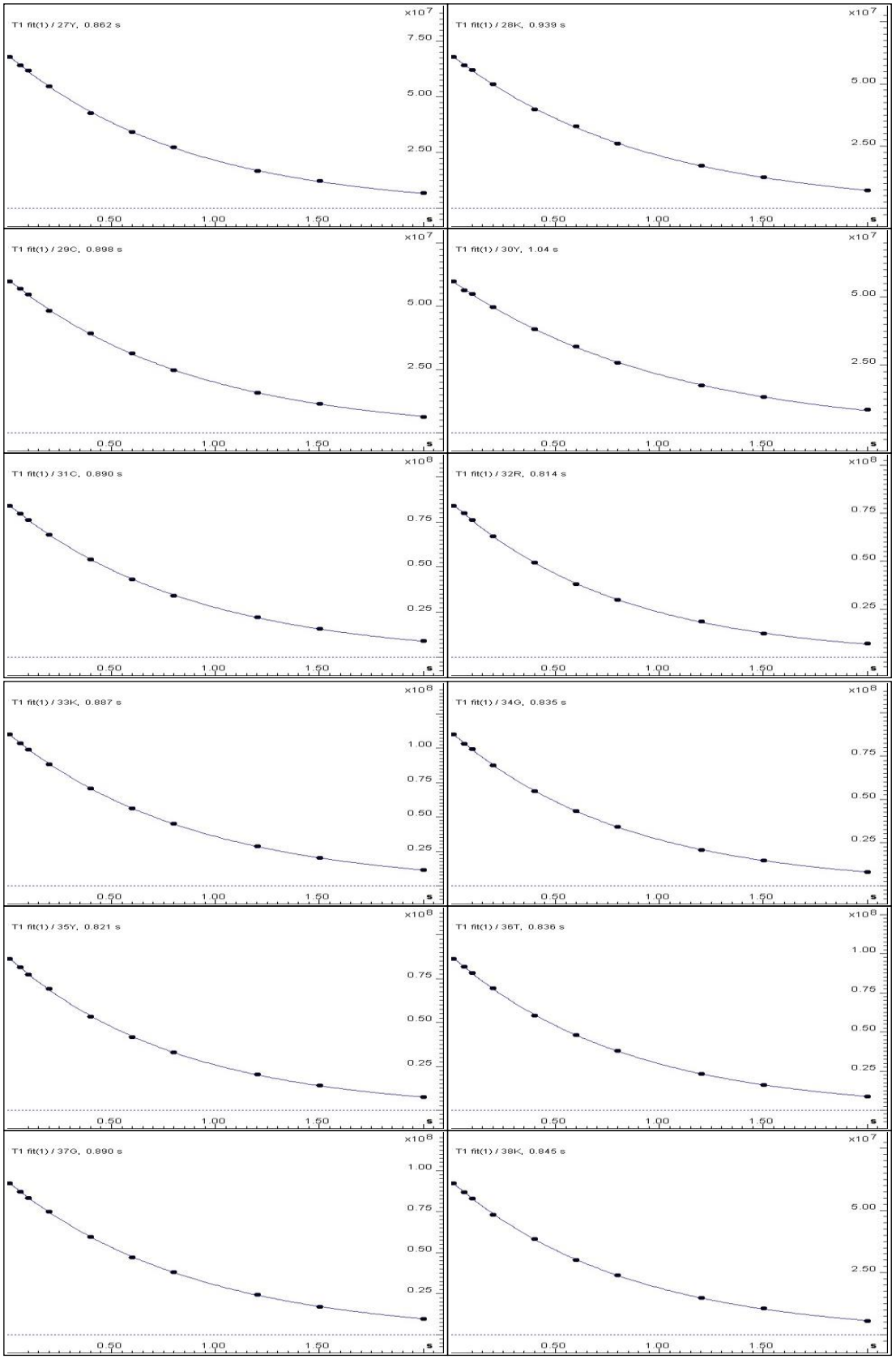
30Y	125.032	9.357	5.62e+07	3.929e+05	1.04	0.01907	Done
31C	123.595	8.788	8.52e+07	2.149e+05	0.890	0.005703	Done
32R	121.032	8.418	8.06e+07	4.574e+05	0.814	0.01157	Done
33K	121.407	8.278	1.11e+08	3.060e+05	0.887	0.006194	Done
34G	112.531	8.319	8.88e+07	2.963e+05	0.835	0.007010	Done
35Y	117.406	7.667	8.77e+07	3.941e+05	0.821	0.009249	Done
36T	112.656	8.923	9.85e+07	4.956e+05	0.836	0.01058	Done
37G	106.780	8.037	9.35e+07	3.475e+05	0.890	0.008398	Done
38K	124.282	9.123	6.17e+07	2.179e+05	0.845	0.007518	Done
39N	111.656	8.371	5.97e+07	3.629e+05	1.02	0.01611	Done
40C	113.531	7.920	6.79e+07	2.602e+05	0.730	0.006917	Done
41Q	117.344	9.639	5.31e+07	2.714e+05	0.808	0.01033	Done
42Y	123.595	8.647	8.55e+07	4.528e+05	0.864	0.01156	Done
43N	121.219	8.847	9.05e+07	3.792e+05	0.910	0.009717	Done

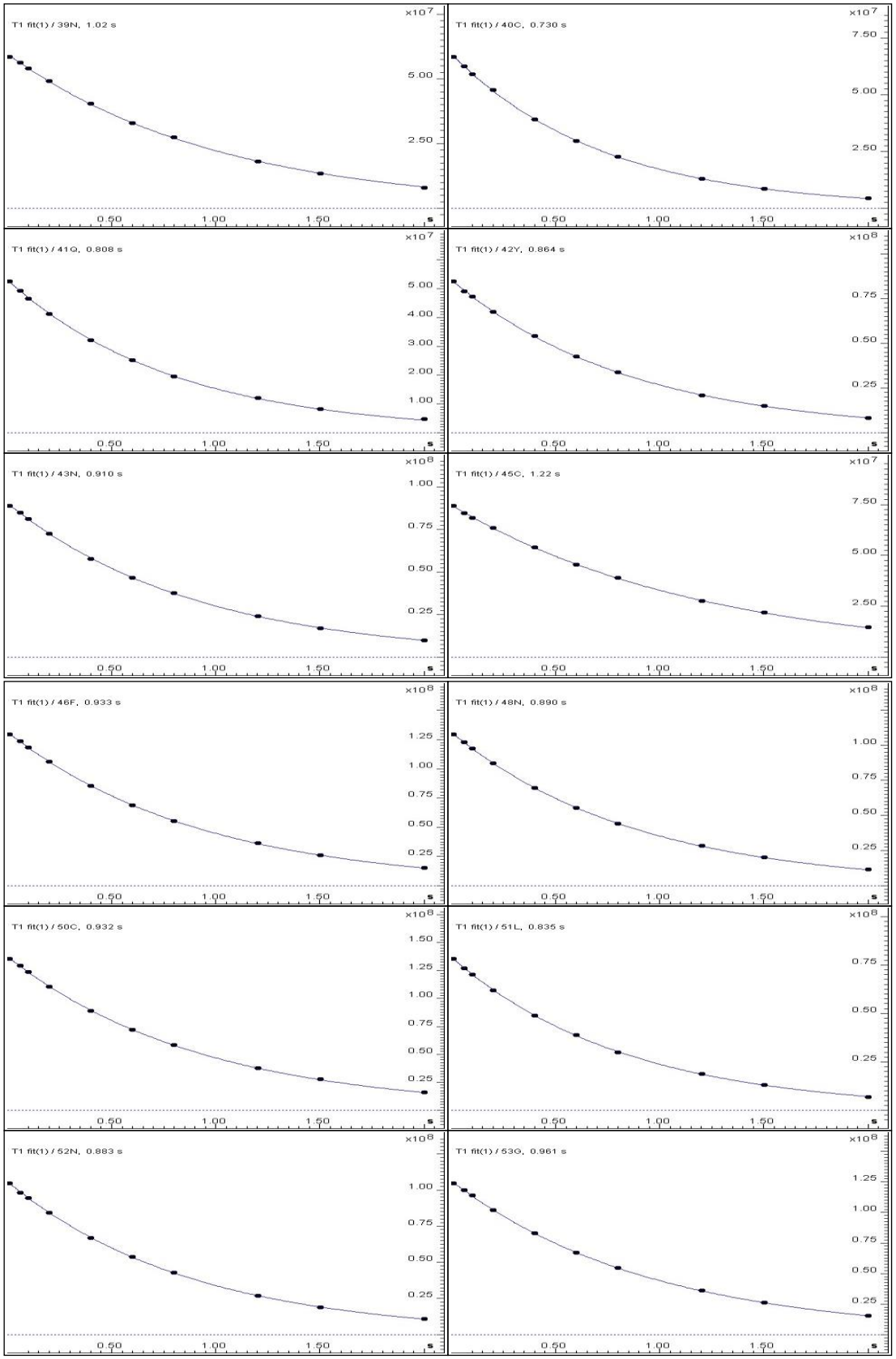
Peak name	F1 [ppm]	F2 [ppm]	Io	error	T1 [s]	error	fitInfo
45C	107.592	7.967	7.48e+07	3.541e+05	1.22	0.01584	Done
46F	121.344	7.345	1.32e+08	2.984e+05	0.933	0.005410	Done
48N	122.157	8.501	1.09e+08	2.886e+05	0.890	0.005972	Done
50C	118.719	7.679	1.38e+08	3.654e+05	0.932	0.006336	Done
51L	125.345	8.395	7.92e+07	2.851e+05	0.835	0.007557	Done
52N	110.655	8.512	1.06e+08	3.594e+05	0.883	0.007606	Done
53G	102.404	8.172	1.26e+08	3.828e+05	0.961	0.007528	Done
54G	106.842	7.667	1.29e+08	4.373e+05	0.764	0.006440	Done
55T	115.719	8.518	1.04e+08	3.021e+05	0.978	0.007336	Done
56C	128.095	9.011	8.80e+07	2.837e+05	0.834	0.006759	Done
57G	118.031	8.782	1.13e+08	3.360e+05	0.830	0.006191	Done
59V	114.718	7.356	1.05e+08	1.750e+05	0.837	0.003508	Done
61K	121.594	8.213	1.38e+08	3.441e+05	0.900	0.005713	Done
61G	109.468	8.060	1.94e+08	5.594e+05	0.681	0.004839	Done
62Y	123.657	8.301	1.46e+08	3.975e+05	0.903	0.006270	Done
64Y	118.594	8.735	1.04e+08	4.544e+05	0.766	0.008325	Done
65Y	114.406	6.822	1.41e+08	3.550e+05	0.764	0.004785	Done
66K	118.844	8.882	9.35e+07	2.866e+05	0.846	0.006528	Done

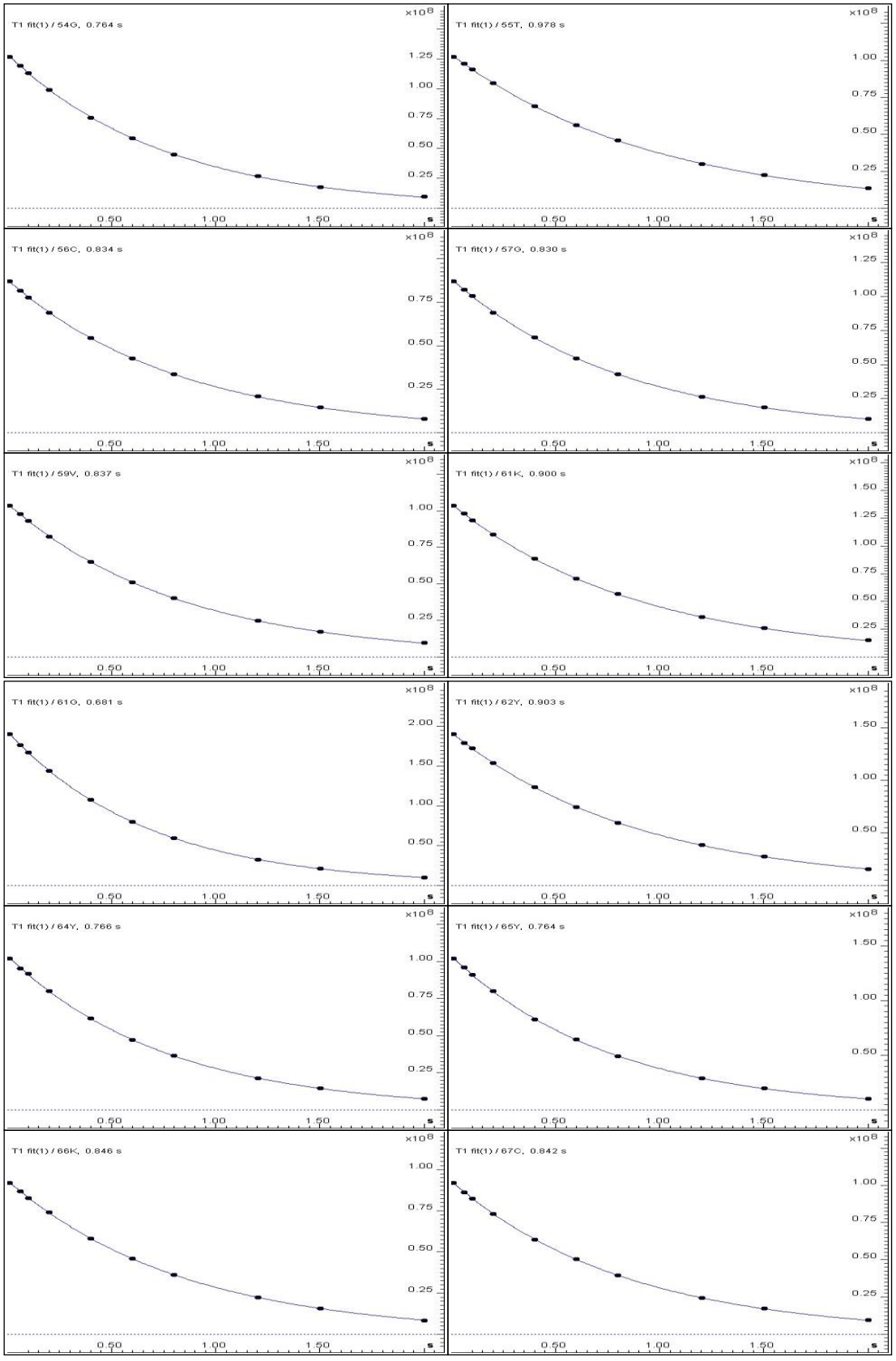
67C	123.220	9.246	1.03e+08	3.262e+05	0.842	0.006723	Done
68S	120.907	9.164	9.87e+07	3.302e+05	0.990	0.008579	Done
69C	124.907	8.999	1.25e+08	3.423e+05	0.873	0.006060	Done
71Y	119.219	8.360	1.58e+08	3.853e+05	0.914	0.005675	Done
72G	113.031	8.195	1.24e+08	3.220e+05	0.879	0.005781	Done
73Y	118.594	7.902	1.44e+08	2.455e+05	0.798	0.003395	Done
74Y	119.094	9.351	1.26e+08	4.407e+05	0.839	0.007372	Done
75G	106.717	8.060	1.21e+08	4.243e+05	0.889	0.007935	Done
76K	122.532	9.187	1.05e+08	3.374e+05	0.821	0.006597	Done
77Q	113.343	8.254	1.06e+08	2.936e+05	0.938	0.006659	Done
79Q	117.906	9.639	1.06e+08	4.137e+05	0.878	0.008675	Done
80L	124.345	8.448	1.04e+08	2.703e+05	0.829	0.005418	Done
82K	125.595	8.231	1.38e+08	5.154e+05	0.860	0.008132	Done
83Y	125.595	7.603	2.24e+08	6.378e+05	0.904	0.006550	Done

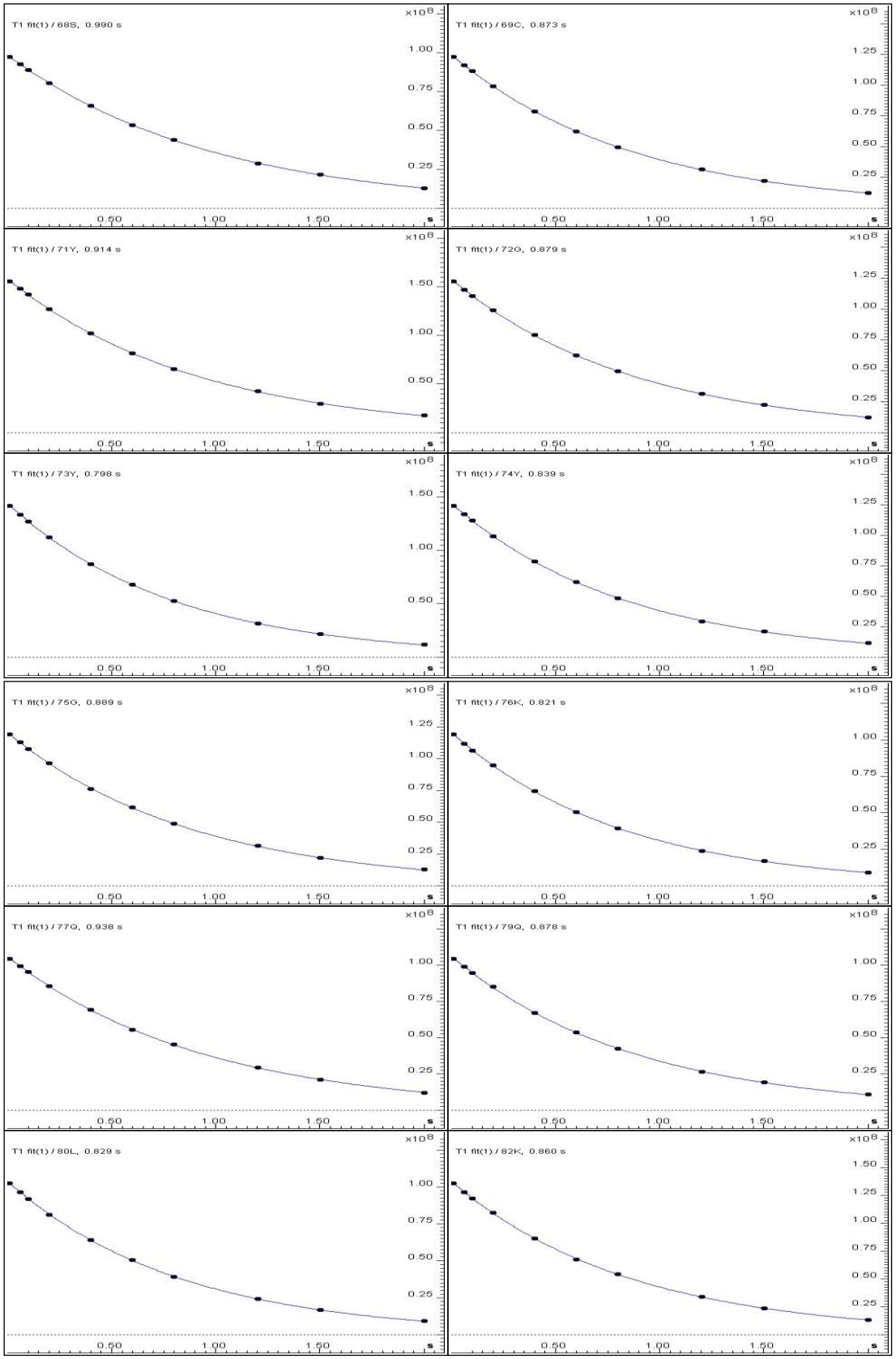


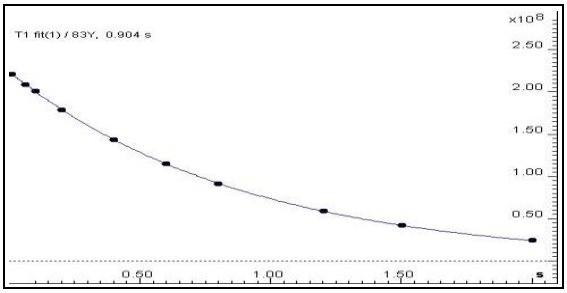

T1 fit(1) / 14R, 0.861 s
T1 fit(1) / 15N, 0.836 s
T1 fit(1) / 16G, 0.840 s
T1 fit(1) / 17G, 0.778 s
T1 fit(1) / 18T, 1.04 s
T1 fit(1) / 19C, 0.842 s
T1 fit(1) / 20K, 0.931 s
T1 fit(1) / 21K, 0.902 s
T1 fit(1) / 23G, 0.932 s
T1 fit(1) / 24L, 0.807 s
T1 fit(1) / 25Y, 0.780 s
T1 fit(1) / 26S, 0.922 s


T1 fit(1) / 27Y, 0.862 s
T1 fit(1) / 28K, 0.939 s
T1 fit(1) / 29C, 0.898 s
T1 fit(1) / 30Y, 1.04 s
T1 fit(1) / 31C, 0.890 s
T1 fit(1) / 32R, 0.814 s
T1 fit(1) / 33K, 0.887 s
T1 fit(1) / 34G, 0.835 s
T1 fit(1) / 35Y, 0.821 s
T1 fit(1) / 36T, 0.836 s
T1 fit(1) / 37G, 0.890 s
T1 fit(1) / 38K, 0.845 s
0.50
1.00
1.50
s


T1 fit(1) / 39N, 1.02 s
T1 fit(1) / 40C, 0.730 s
T1 fit(1) / 41Q, 0.808 s
T1 fit(1) / 42Y, 0.864 s
T1 fit(1) / 43N, 0.910 s
T1 fit(1) / 45C, 1.22 s
T1 fit(1) / 46F, 0.933 s
T1 fit(1) / 48N, 0.890 s
T1 fit(1) / 50C, 0.932 s
T1 fit(1) / 51L, 0.835 s
T1 fit(1) / 52N, 0.883 s
T1 fit(1) / 53G, 0.961 s
0.50
1.00
1.50
s


T1 fit(1) / 54G, 0.764 s
T1 fit(1) / 55T, 0.978 s
T1 fit(1) / 56C, 0.834 s
T1 fit(1) / 57G, 0.830 s
T1 fit(1) / 59V, 0.837 s
T1 fit(1) / 61K, 0.900 s
T1 fit(1) / 61G, 0.681 s
T1 fit(1) / 62Y, 0.903 s
T1 fit(1) / 64Y, 0.766 s
T1 fit(1) / 65Y, 0.764 s
T1 fit(1) / 66K, 0.846 s
T1 fit(1) / 67C, 0.842 s
x10^8
s


T1 fit(1) / 68S, 0.990 s
T1 fit(1) / 69C, 0.873 s
T1 fit(1) / 71Y, 0.914 s
T1 fit(1) / 72G, 0.879 s
T1 fit(1) / 73Y, 0.798 s
T1 fit(1) / 74Y, 0.839 s
T1 fit(1) / 75G, 0.889 s
T1 fit(1) / 76K, 0.821 s
T1 fit(1) / 77Q, 0.938 s
T1 fit(1) / 79Q, 0.878 s
T1 fit(1) / 80L, 0.829 s
T1 fit(1) / 82K, 0.860 s
x10^8
s



Dynamic Center Analysis of T2 at 800 MHz

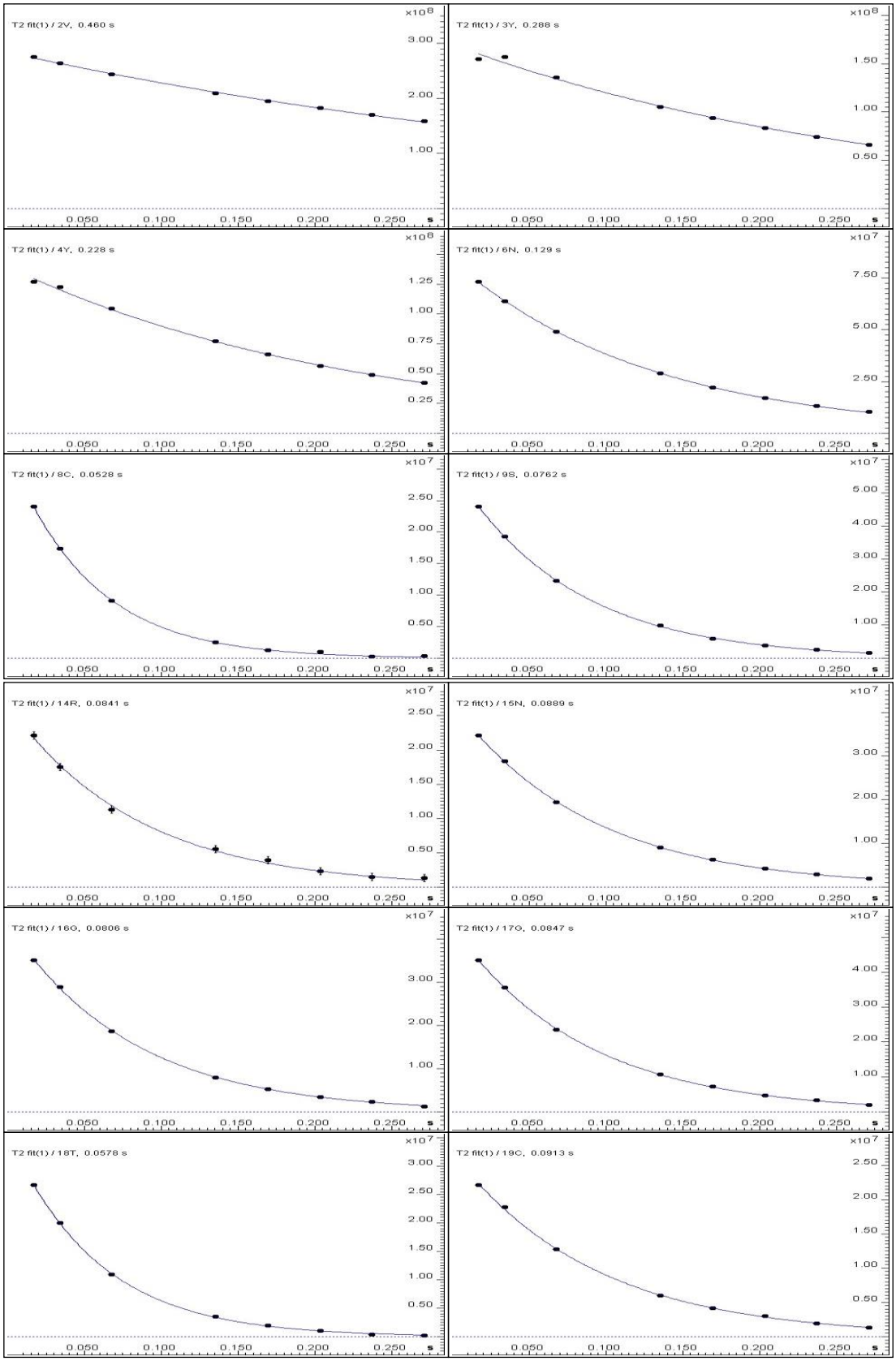
Fitted function:	f(t) = Io * exp (-t/T)
Random error estimation of data:	RMS per spectrum (or trace/plane)
Systematic error estimation of data:	worst case per peak scenario
Fit parameter Error estimation method:	from fit using arbitray y uncertainties
Confidence level:	95%
Used peaks:	peaks from peaklist.xml at spectrum
Used integrals:	peak intensities
Used Time:	integrals related to replicate values replaced by mean

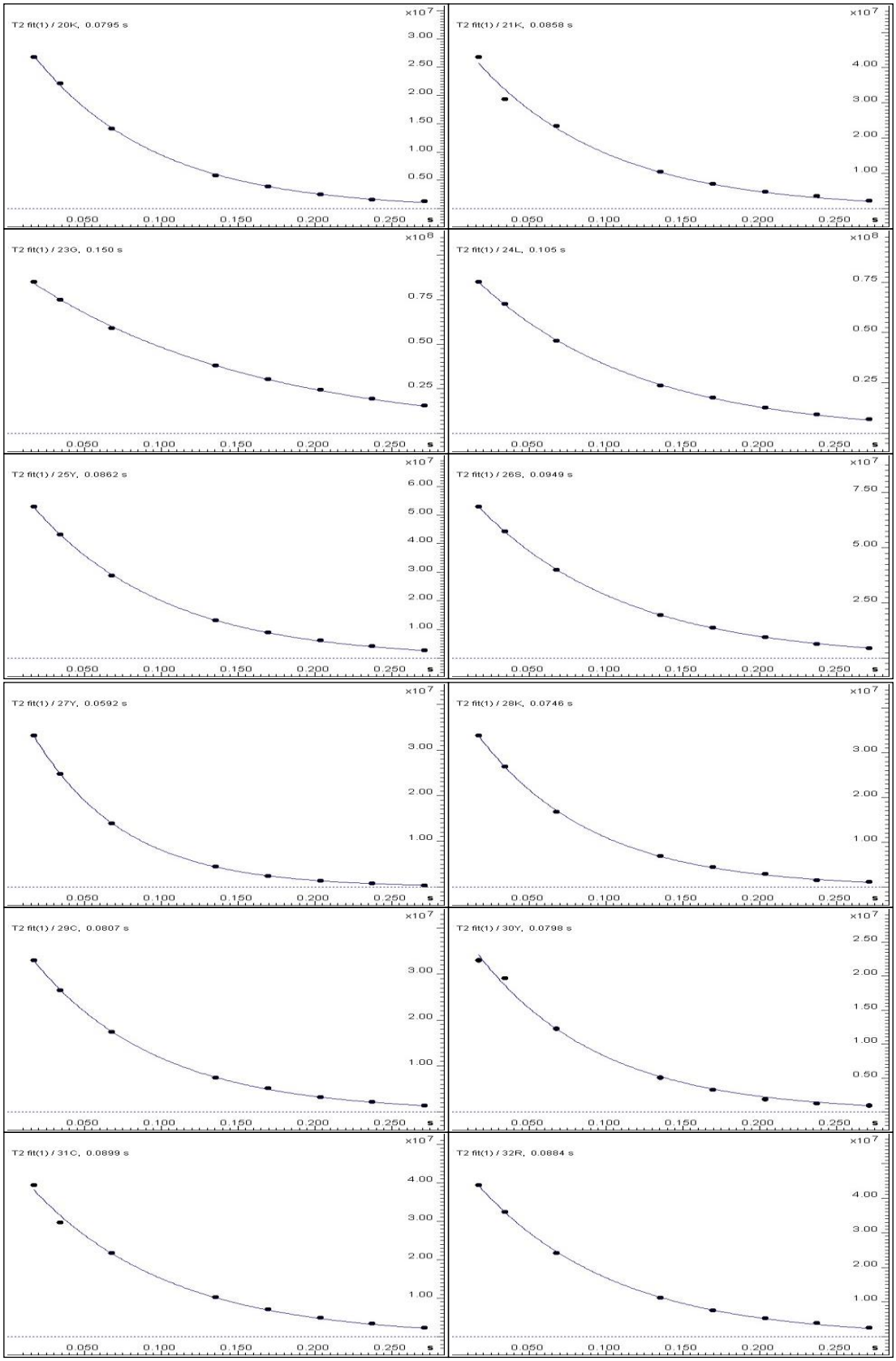
Peak name	F1 [ppm]	F2 [ppm]	Io	error	T2 [s]	error	fitInfo
2V	119.344	8.260	2.85e+08	2.860e+06	0.460	0.01542	Done
3Y	124.720	8.166	1.70e+08	6.761e+06	0.288	0.02734	Done
4Y	125.220	7.890	1.40e+08	3.041e+06	0.228	0.01035	Done
6N	118.344	8.330	8.34e+07	5.161e+05	0.129	0.001376	Done
8C	112.281	7.767	3.32e+07	7.207e+05	0.0528	0.001798	Done
9S	116.094	7.374	5.74e+07	4.378e+05	0.0762	0.0009665	Done
14R	122.845	8.512	2.66e+07	1.174e+06	0.0841	0.006191	Done
15N	114.031	9.076	4.22e+07	3.227e+05	0.0889	0.001140	Done
16G	106.968	8.765	4.36e+07	5.661e+05	0.0806	0.001744	Done
17G	105.655	7.362	5.32e+07	4.344e+05	0.0847	0.001156	Done
18T	119.157	8.841	3.59e+07	4.690e+05	0.0578	0.001211	Done
19C	127.720	8.424	2.70e+07	5.105e+05	0.0913	0.002891	Done
20K	130.346	9.246	3.35e+07	6.508e+05	0.0795	0.002572	Done
21K	123.907	8.753	5.04e+07	4.167e+06	0.0858	0.01186	Done

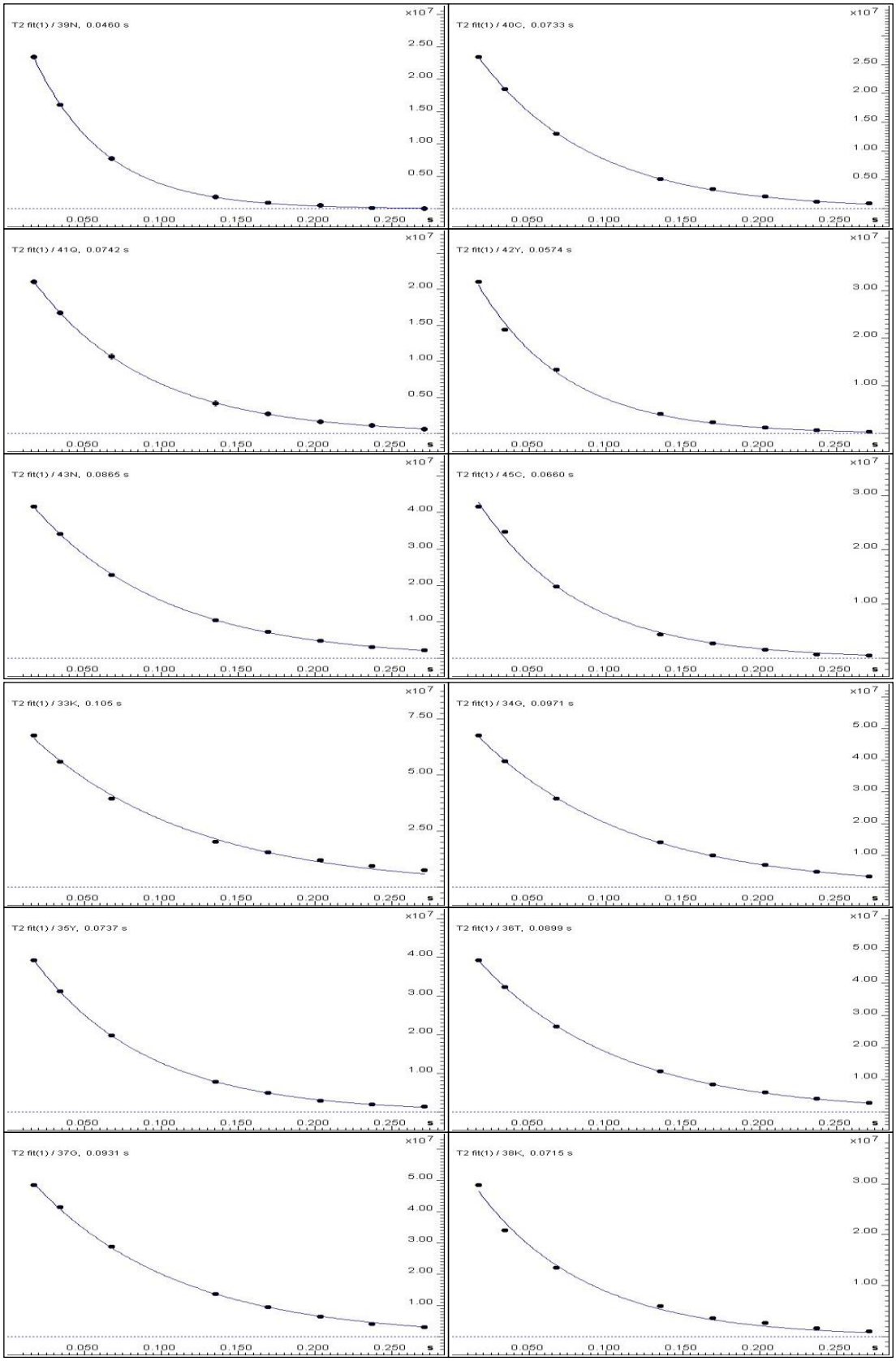
23G	111.468	8.148	9.46e+07	1.302e+06	0.150	0.003663	Done
24L	119.532	8.154	8.85e+07	9.389e+05	0.105	0.001880	Done
25Y	111.656	8.014	6.43e+07	7.409e+05	0.0862	0.001662	Done
26S	113.656	7.186	8.23e+07	2.650e+05	0.0949	0.0005132	Done
27Y	117.531	7.931	4.42e+07	3.563e+05	0.0592	0.0007694	Done
28K	120.844	9.070	4.25e+07	5.506e+05	0.0746	0.001606	Done
29C	121.657	9.046	4.07e+07	4.411e+05	0.0807	0.001460	Done
30Y	125.032	9.357	2.87e+07	1.739e+06	0.0798	0.008073	Done
31C	123.532	8.788	4.61e+07	2.767e+06	0.0899	0.009030	Done
32R	121.032	8.418	5.30e+07	7.546e+05	0.0884	0.002105	Done
33K	121.344	8.278	7.81e+07	3.685e+06	0.105	0.008389	Done
34G	112.531	8.319	5.68e+07	6.252e+05	0.0971	0.001797	Done
35Y	117.406	7.673	4.95e+07	3.512e+05	0.0737	0.0008684	Done
36T	112.656	8.923	5.69e+07	3.729e+05	0.0899	0.0009878	Done
37G	106.780	8.037	5.90e+07	1.098e+06	0.0931	0.002905	Done
38K	124.157	9.123	3.62e+07	3.327e+06	0.0715	0.01090	Done
39N	111.718	8.371	3.39e+07	5.877e+05	0.0460	0.001196	Done
40C	113.593	7.920	3.32e+07	3.178e+05	0.0733	0.001164	Done
41Q	117.344	9.639	2.65e+07	2.081e+05	0.0742	0.0009677	Done
42Y	123.532	8.647	4.19e+07	2.853e+06	0.0574	0.006260	Done
43N	121.219	8.847	5.07e+07	3.848e+05	0.0865	0.001098	Done

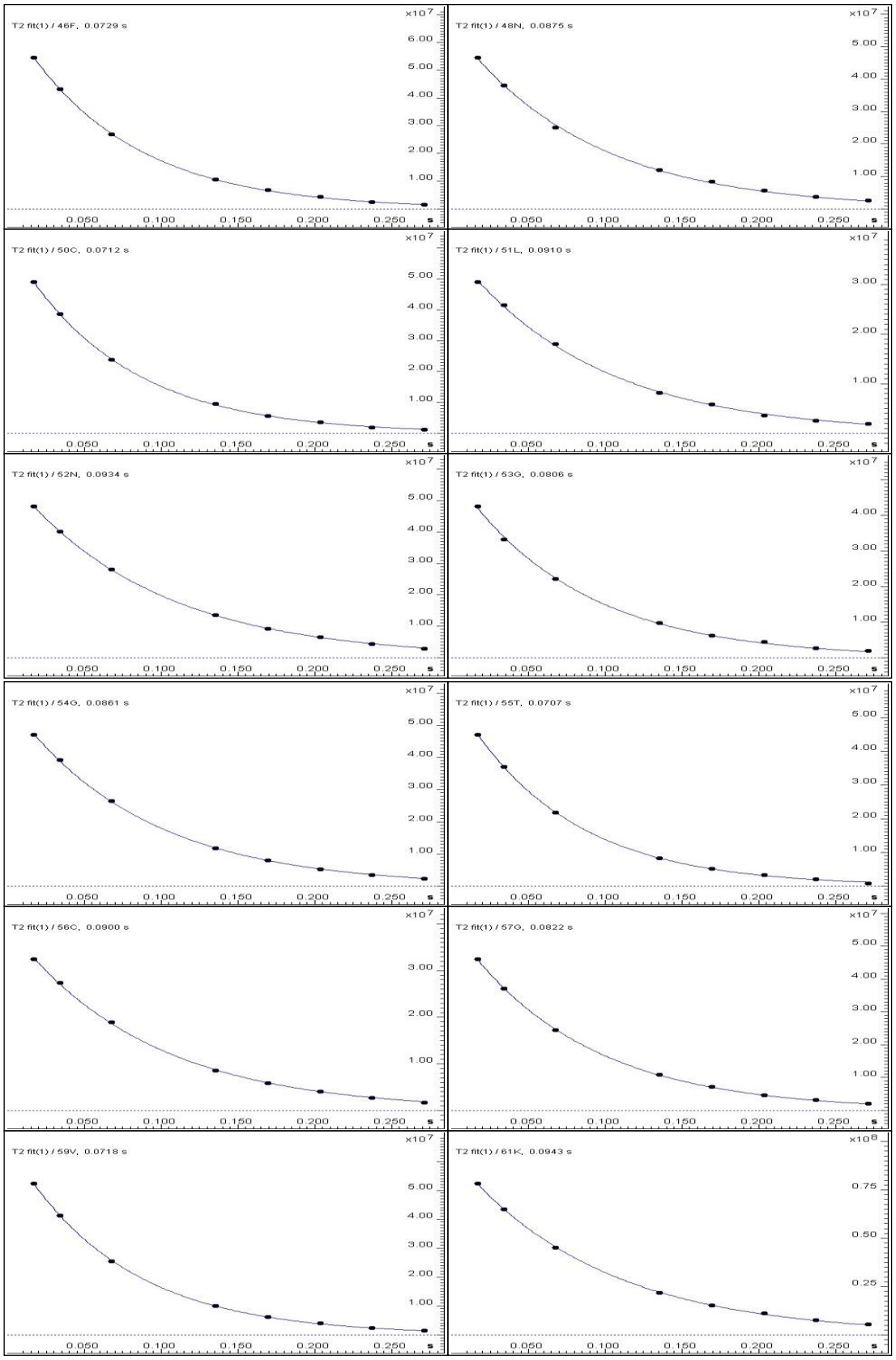
Peak name	F1 [ppm]	F2 [ppm]	Io	error	T2 [s]	error	fitInfo
45C	107.655	7.967	3.72e+07	2.152e+06	0.0660	0.006262	Done
46F	121.344	7.345	6.89e+07	4.293e+05	0.0729	0.0007532	Done
48N	122.220	8.501	5.64e+07	1.138e+06	0.0875	0.002958	Done
50C	118.657	7.679	6.23e+07	6.026e+05	0.0712	0.001141	Done
51L	125.345	8.395	3.72e+07	8.091e+05	0.0910	0.003319	Done
52N	110.718	8.512	5.79e+07	3.959e+05	0.0934	0.001071	Done
53G	102.467	8.172	5.19e+07	1.350e+06	0.0806	0.003495	Done
54G	106.843	7.667	5.77e+07	6.162e+05	0.0861	0.001537	Done
55T	115.719	8.518	5.72e+07	5.903e+05	0.0707	0.001206	Done
56C	128.095	9.011	3.96e+07	7.030e+05	0.0900	0.002673	Done
57G	118.031	8.788	5.65e+07	5.500e+05	0.0822	0.001336	Done

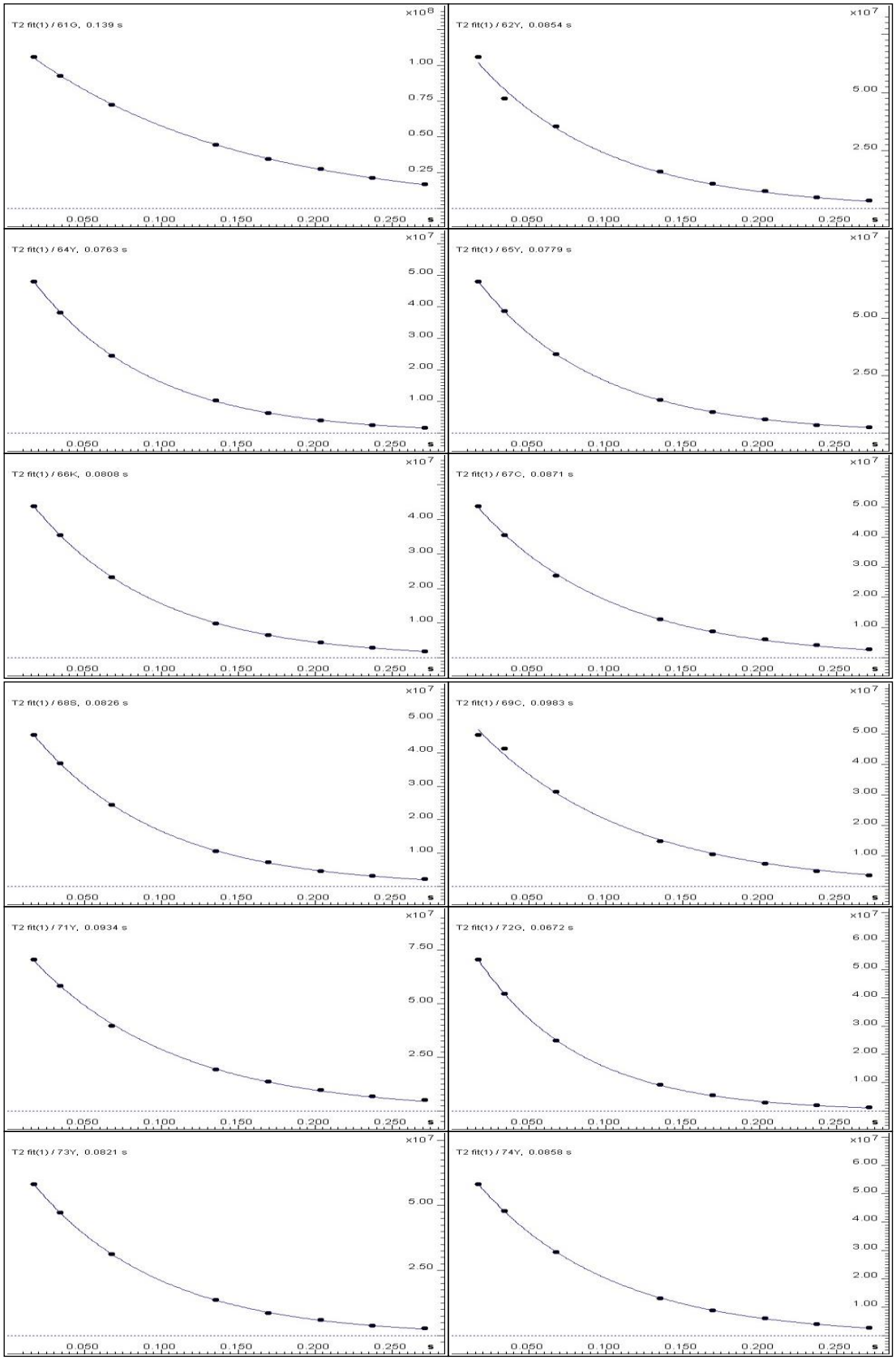
59V	114.719	7.356	6.63e+07	5.272e+05	0.0718	0.0009463	Done
61K	121.594	8.213	9.36e+07	9.673e+05	0.0943	0.001636	Done
61G	109.468	8.060	1.19e+08	1.078e+06	0.139	0.002189	Done
62Y	123.595	8.301	7.67e+07	6.053e+06	0.0854	0.01126	Done
64Y	118.594	8.735	6.01e+07	4.283e+05	0.0763	0.0009046	Done
65Y	114.406	6.822	8.24e+07	4.293e+05	0.0779	0.0006760	Done
66K	118.782	8.882	5.41e+07	2.676e+05	0.0808	0.0006669	Done
67C	123.157	9.246	6.07e+07	1.138e+06	0.0871	0.002730	Done
68S	120.907	9.164	5.58e+07	3.642e+05	0.0826	0.0009010	Done
69C	124.907	8.999	6.13e+07	3.131e+06	0.0983	0.008440	Done
71Y	119.219	8.360	8.42e+07	1.599e+06	0.0934	0.002974	Done
72G	113.031	8.201	6.89e+07	7.127e+05	0.0672	0.001144	Done
73Y	118.594	7.902	7.16e+07	4.621e+05	0.0821	0.0008841	Done
74Y	119.094	9.357	6.53e+07	3.010e+05	0.0858	0.0006616	Done
75G	106.717	8.055	6.70e+07	1.164e+06	0.0906	0.002635	Done
76K	122.532	9.181	6.81e+07	6.983e+05	0.0841	0.001440	Done
77Q	113.343	8.260	5.85e+07	7.272e+05	0.0658	0.001341	Done
79Q	117.906	9.639	5.51e+07	5.349e+05	0.0713	0.001145	Done
80L	124.282	8.454	5.12e+07	2.238e+06	0.0631	0.004505	Done
82K	125.595	8.236	8.94e+07	9.894e+05	0.103	0.001919	Done
83Y	125.595	7.603	1.52e+08	8.155e+05	0.141	0.001325	Done

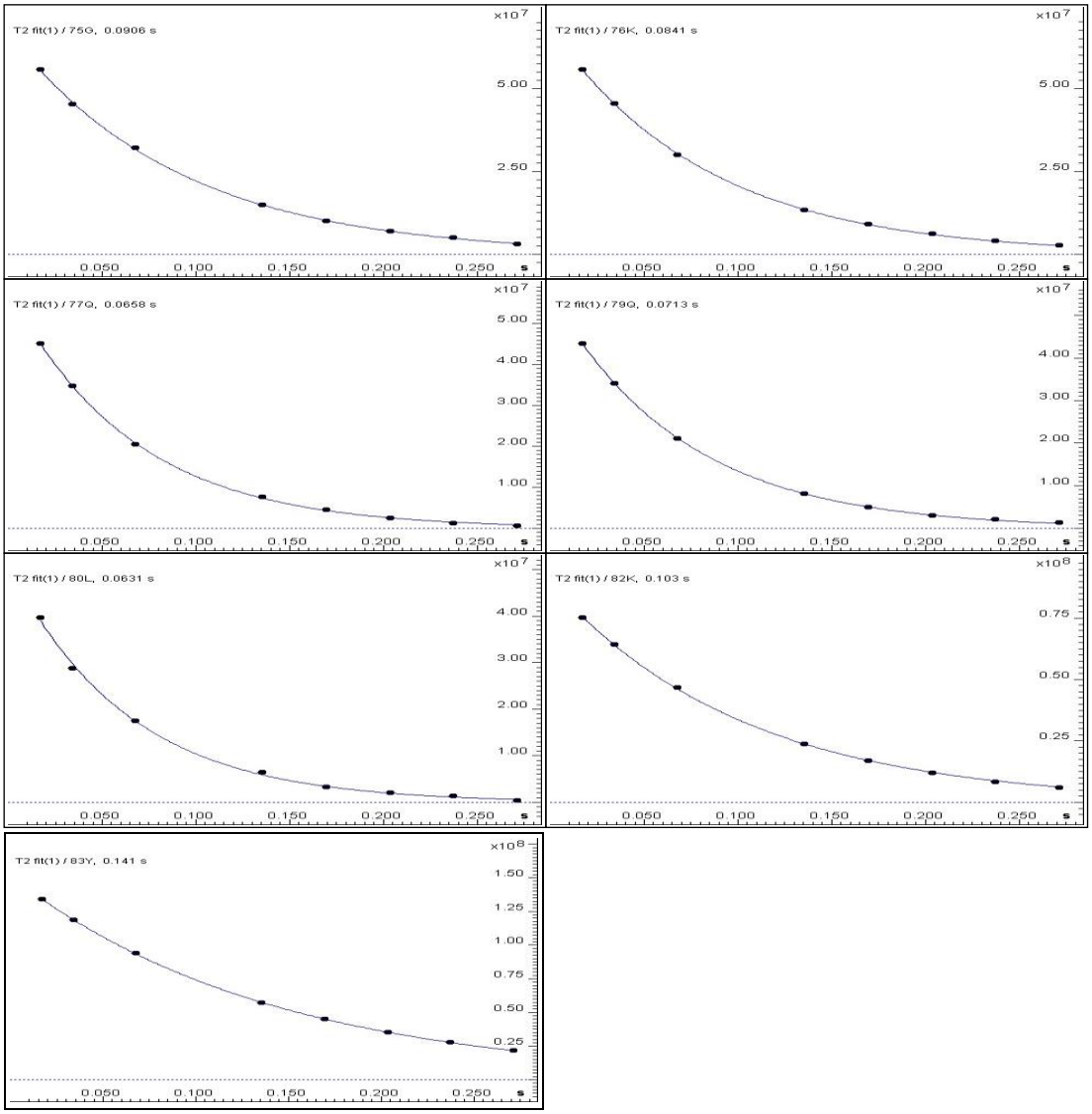

T2 fit(1) / 2V, 0.460 s
T2 fit(1) / 3Y, 0.288 s
T2 fit(1) / 4Y, 0.228 s
T2 fit(1) / 6N, 0.129 s
T2 fit(1) / 8C, 0.0528 s
T2 fit(1) / 9S, 0.0762 s
T2 fit(1) / 14R, 0.0841 s
T2 fit(1) / 15N, 0.0889 s
T2 fit(1) / 16G, 0.0806 s
T2 fit(1) / 17G, 0.0847 s
T2 fit(1) / 18T, 0.0578 s
T2 fit(1) / 19C, 0.0913 s


T2 fit(1) / 20K, 0.0795 s
T2 fit(1) / 21K, 0.0858 s
T2 fit(1) / 23G, 0.150 s
T2 fit(1) / 24L, 0.105 s
T2 fit(1) / 25Y, 0.0862 s
T2 fit(1) / 26S, 0.0949 s
T2 fit(1) / 27Y, 0.0592 s
T2 fit(1) / 28K, 0.0746 s
T2 fit(1) / 29C, 0.0807 s
T2 fit(1) / 30Y, 0.0798 s
T2 fit(1) / 31C, 0.0899 s
T2 fit(1) / 32R, 0.0884 s
0.050
0.100
0.150
0.200
0.250
s


T2 fit(1) / 39N, 0.0460 s
T2 fit(1) / 40C, 0.0733 s
T2 fit(1) / 41Q, 0.0742 s
T2 fit(1) / 42Y, 0.0574 s
T2 fit(1) / 43N, 0.0865 s
T2 fit(1) / 45C, 0.0660 s
T2 fit(1) / 33K, 0.105 s
T2 fit(1) / 34G, 0.0971 s
T2 fit(1) / 35Y, 0.0737 s
T2 fit(1) / 36T, 0.0899 s
T2 fit(1) / 37G, 0.0931 s
T2 fit(1) / 38K, 0.0715 s
x10^7
s


T2 fit(1) / 46F, 0.0729 s
T2 fit(1) / 48N, 0.0875 s
T2 fit(1) / 50C, 0.0712 s
T2 fit(1) / 51L, 0.0910 s
T2 fit(1) / 52N, 0.0934 s
T2 fit(1) / 53G, 0.0806 s
T2 fit(1) / 54G, 0.0861 s
T2 fit(1) / 55T, 0.0707 s
T2 fit(1) / 56C, 0.0900 s
T2 fit(1) / 57G, 0.0822 s
T2 fit(1) / 59V, 0.0718 s
T2 fit(1) / 61K, 0.0943 s


T2 fit(1) / 61G, 0.139 s
T2 fit(1) / 62Y, 0.0854 s
T2 fit(1) / 64Y, 0.0763 s
T2 fit(1) / 65Y, 0.0779 s
T2 fit(1) / 66K, 0.0808 s
T2 fit(1) / 67C, 0.0871 s
T2 fit(1) / 68S, 0.0826 s
T2 fit(1) / 69C, 0.0983 s
T2 fit(1) / 71Y, 0.0934 s
T2 fit(1) / 72G, 0.0672 s
T2 fit(1) / 73Y, 0.0821 s
T2 fit(1) / 74Y, 0.0858 s



Dynamic Center Analysis of NOE at 800 MHz

Random error estimation of data:	RMS per spectrum (or trace/plane)
Systematic error estimation of data:	none
Used peaks:	peaks from peaklist.xml at spectrum
Used integrals:	peak intensities

Peak name	F1 [ppm]	F2 [ppm]	NOE [ ]	error	fitInfo
2V	119.391	8.260	-0.289	0.0001009	Fail
3Y	124.767	8.166	-0.274	0.0001402	Fail

4Y	125.267	7.890	0.0484	0.0001486	Fail
6N	118.391	8.330	0.493	0.0002969	Fail
8C	112.265	7.761	0.716	0.0007086	Fail
9S	116.078	7.368	0.773	0.0004845	Fail
14R	122.829	8.512	0.716	0.0007844	Fail
15N	114.015	9.076	0.824	0.0006084	Fail
16G	106.952	8.765	0.856	0.0005710	Fail
17G	105.639	7.356	0.782	0.0004983	Fail
18T	119.204	8.841	0.736	0.0006898	Fail
19C	127.767	8.424	0.781	0.0008070	Fail
20K	130.393	9.240	0.781	0.0006980	Fail
21K	124.017	8.753	0.760	0.0006022	Fail
23G	111.453	8.143	0.302	0.0002333	Fail
24L	119.516	8.154	0.605	0.0003585	Fail
25Y	111.578	8.008	0.602	0.0004526	Fail
26S	113.703	7.180	0.577	0.0003407	Fail
27Y	117.516	7.931	0.744	0.0006441	Fail
28K	120.891	9.064	0.728	0.0006523	Fail
29C	121.704	9.046	0.787	0.0006838	Fail
30Y	125.079	9.357	0.773	0.0007236	Fail
31C	123.579	8.788	0.785	0.0005293	Fail
32R	121.016	8.418	0.836	0.0005339	Fail
33K	121.391	8.278	0.725	0.0003781	Fail
34G	112.515	8.319	0.737	0.0004922	Fail
35Y	117.391	7.667	0.826	0.0005573	Fail
36T	112.640	8.923	0.813	0.0004517	Fail
37G	106.796	8.037	0.768	0.0004579	Fail
38K	124.267	9.123	0.805	0.0006815	Fail
39N	111.640	8.371	0.780	0.0007466	Fail
40C	113.578	7.920	0.801	0.0006590	Fail
41Q	117.360	9.639	0.849	0.0008305	Fail
42Y	123.579	8.647	0.807	0.0005622	Fail
43N	121.204	8.847	0.768	0.0004749	Fail
45C	107.608	7.967	0.769	0.0005172	Fail

46F	121.391	7.345	0.830	0.0003549	Fail
48N	122.204	8.501	0.838	0.0004551	Fail
50C	118.703	7.679	0.814	0.0003605	Fail

Peak name	F1 [ppm]	F2 [ppm]	NOE [ ]	error	fitInfo
51L	125.329	8.395	0.764	0.0004958	Fail
52N	110.702	8.512	0.788	0.0003961	Fail
53G	102.451	8.172	0.799	0.0003473	Fail
54G	106.827	7.661	0.790	0.0003295	Fail
55T	115.703	8.518	0.814	0.0004585	Fail
56C	128.080	9.011	0.810	0.0004944	Fail
57G	118.078	8.782	0.776	0.0004028	Fail
59V	114.765	7.356	0.764	0.0004087	Fail
61K	121.641	8.213	0.735	0.0003097	Fail
61G	109.452	8.060	0.635	0.0002259	Fail
62Y	123.704	8.301	0.757	0.0003029	Fail
64Y	118.641	8.735	0.755	0.0004294	Fail
65Y	114.390	6.816	0.785	0.0003422	Fail
66K	118.828	8.882	0.778	0.0004824	Fail
67C	123.204	9.246	0.792	0.0004508	Fail
68S	120.891	9.164	0.754	0.0004248	Fail
69C	124.954	8.999	0.789	0.0003465	Fail
71Y	119.204	8.360	0.730	0.0002829	Fail
72G	113.015	8.195	0.697	0.0003445	Fail
73Y	118.578	7.902	0.786	0.0003637	Fail
74Y	119.141	9.351	0.765	0.0003512	Fail
75G	106.702	8.060	0.759	0.0003568	Fail
76K	122.516	9.187	0.754	0.0004110	Fail
77Q	113.328	8.254	0.777	0.0004321	Fail
79Q	117.891	9.639	0.788	0.0004511	Fail
80L	124.329	8.448	0.801	0.0004378	Fail
82K	125.579	8.231	0.613	0.0002807	Fail
83Y	125.579	7.603	0.392	0.0001609	Fail

References

1. Ikada Y. Challenges in tissue engineering. *J R Soc Interface*. 2006;3(10):589-601. doi:10.1098/rsif.2006.0124
2. Rusk RD. Science. *Educ Forum*. 1950;15(1):119-120. doi:10.1080/00131725009342110
3. The Edwin Smith Surgical Papyrus, Volume 1: Hieroglyphic Transliteration, Translation, and Commentary | The Oriental Institute of the University of Chicago. Accessed December 29, 2020. https://oi.uchicago.edu/research/publications/oip/edwin-smith-surgical-papyrus-volume-1-hieroglyphic-transliteration
4. Sundaram C, Keenan A. Evolution of hemostatic agents in surgical practice. *Indian J Urol*. 2010;26(3):374. doi:10.4103/0970-1591.70574
5. Liu D, Nikoo M, Boran G, Zhou P, Regenstein JM. Collagen and gelatin. *Annu Rev Food Sci Technol*. 2015;6:527-557. doi:10.1146/annurev-food-031414-111800
6. Spotnitz WD, Burks S. Hemostats, sealants, and adhesives: Components of the surgical toolbox. *Transfusion*. 2008;48(7):1502-1516. doi:10.1111/j.1537-2995.2008.01703.x
7. Meyer M. Processing of collagen based biomaterials and the resulting materials properties. *Biomed Eng Online*. 2019;18(1):1-74. doi:10.1186/s12938-019-0647-0
8. Chi JH. MeTro : A New. 2018;82(3):2017-2018.
9. Annabi N, Rana D, Shirzaei Sani E, et al. Engineering a sprayable and elastic hydrogel adhesive with antimicrobial properties for wound healing. *Biomaterials*. 2017;139:229-243. doi:10.1016/j.biomaterials.2017.05.011
10. Coover HW. Discovery of superglue shows power of pursuing the unexplained. *Res Technol Manag*. 2000;43(5):36-39. doi:10.1080/08956308.2000.11671379
11. Donkerwolcke M, Burny F, Muster D. Tissues and bone adhesives historical aspects. *Biomaterials*. 1998;19(16):1461-1466. doi:10.1016/S0142-9612(98)00059-3
12. Tissue Adhesive for the Topical Approximation of Skin - Class II Special Controls Guidance for Industry and FDA Staff | FDA. Accessed January 5, 2021. https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-

radiation-emitting-products/tissue-adhesive-topical-approximation-skin-class-ii-special-controls-guidance-industry-and-fda-staff

13. Zhu W, Chuah YJ, Wang DA. Bioadhesives for internal medical applications: A review. *Acta Biomater*. 2018;74:1-16. doi:10.1016/j.actbio.2018.04.034
14. Mehdizadeh M, Yang J. Design Strategies and Applications of Tissue Bioadhesives. *Macromol Biosci*. 2013;13(3):271-288. doi:10.1002/mabi.201200332
15. Shen T, Sun Y, Sun C, Lu M. Preparation and characterization of polyurethane bioadhesive from Hydroxyl-terminated polylactide and imidazole-blocked isocyanate. *Polym*. 2013;37(2):232-239. doi:10.7317/pk.2013.37.2.232
16. Lin CC, Anseth KS. PEG hydrogels for the controlled release of biomolecules in regenerative medicine. *Pharm Res*. 2009;26(3):631-643. doi:10.1007/s11095-008-9801-2
17. Turecek PL, Siekmann J. PEG-protein conjugates: Nonclinical and clinical toxicity considerations. In: *Polymer-Protein Conjugates: From Pegylation and Beyond*. Elsevier; 2019:61-101. doi:10.1016/B978-0-444-64081-9.00004-8
18. Pinnaratip R, Bhuiyan MSA, Meyers K, Rajachar RM, Lee BP. Multifunctional Biomedical Adhesives. *Adv Healthc Mater*. 2019;8(11):1-31. doi:10.1002/adhm.201801568
19. Lefevre M, Flammang P, Aranko AS, et al. Sea star-inspired recombinant adhesive proteins self-assemble and adsorb on surfaces in aqueous environments to form cytocompatible coatings. *Acta Biomater*. 2020;112(xxxx):62-74. doi:10.1016/j.actbio.2020.05.036
20. Jo YK, Kim HJ, Jeong Y, Joo K Il, Cha HJ. Biomimetic Surface Engineering of Biomaterials by Using Recombinant Mussel Adhesive Proteins. *Adv Mater Interfaces*. 2018;5(9):1-13. doi:10.1002/admi.201800068
21. Hennebert E, Wattiez R, Demeuldre M, et al. Sea star tenacity mediated by a protein that fragments, then aggregates. *Proc Natl Acad Sci U S A*. 2014;111(17):6317-6322. doi:10.1073/pnas.1400089111
22. Xu X, Chen X, Li J. Natural protein bioinspired materials for regeneration of hard tissues. *J Mater Chem B*. 2020;8(11):2199-2215. doi:10.1039/d0tb00139b
23. Kord Forooshani P, Lee BP. Recent approaches in designing bioadhesive materials

inspired by mussel adhesive protein. *J Polym Sci Part A Polym Chem*. 2017;55(1):9-33. doi:10.1002/pola.28368

24. Rahimnejad M, Zhong W. Mussel-inspired hydrogel tissue adhesives for wound closure. *RSC Adv*. 2017;7(75):47380-47396. doi:10.1039/c7ra06743g
25. Ahn BK. Perspectives on Mussel-Inspired Wet Adhesion. *J Am Chem Soc*. 2017;139(30):10166-10171. doi:10.1021/jacs.6b13149
26. Tamarin A, Keller PJ. An L / Itrastructutal Study of the Byssal Thread Forming. 1972;416:401-416.
27. Mediodia DP, DE LEON SMS, Anasco NC, Baylon CC. Shell morphology and anatomy of the Philippine Charru Mussel Mytella charruana (D'orbigny 1842). *Asian Fish Sci*. 2017;30(3):185-194.
28. Eggermont M, Cornillie P, Dierick M, et al. The blue mussel inside: 3D visualization and description of the vascular-related anatomy of Mytilus edulis to unravel hemolymph extraction. *Sci Rep*. 2020;10(1):1-16. doi:10.1038/s41598-020-62933-9
29. Waite JH. Mussel adhesion - Essential footwork. *J Exp Biol*. 2017;220(4):517-530. doi:10.1242/jeb.134056
30. Priemel T, Degtyar E, Dean MN, Harrington MJ. Rapid self-assembly of complex biomolecular architectures during mussel byssus biofabrication. *Nat Commun*. 2017;8(March):1-12. doi:10.1038/ncomms14539
31. Sáez C, Pardo J, Gutierrez E, Brito M, Burzio LO. Immunological studies of the polyphenolic proteins of mussels. *Comp Biochem Physiol -- Part B Biochem*. 1991;98(4):569-572. doi:10.1016/0305-0491(91)90255-C
32. Yu M, Hwang J, Deming TJ. Role of 1-3,4-dihydroxyphenylalanine in mussel adhesive proteins. *J Am Chem Soc*. 1999;121(24):5825-5826. doi:10.1021/ja990469y
33. Barbara S. Adhesion a. *Biochemistry*. 2002;1180:1172-1180.
34. Wei W, Yu J, Broomell C, Israelachvili JN, Waite JH. Hydrophobic enhancement of dopa-mediated adhesion in a mussel foot protein. *J Am Chem Soc*. 2013;135(1):377-383. doi:10.1021/ja309590f
35. Gebbie MA, Wei W, Schrader AM, et al. Tuning underwater adhesion with cation-φ interactions. *Nat Chem*. 2017;9(5):473-479. doi:10.1038/nchem.2720

36. Zhang W, Yang H, Liu F, et al. Molecular interactions between DOPA and surfaces with different functional groups: A chemical force microscopy study. *RSC Adv*. 2017;7(52):32518-32527. doi:10.1039/c7ra04228k
37. Monahan J, Wilker JJ. Specificity of metal ion cross-linking in marine mussel adhesives. *Chem Commun*. 2003;3(14):1672-1673. doi:10.1039/b301052j
38. Sever MJ, Weisser JT, Monahan J, Srinivasan S, Wilker JJ. Metal-Mediated Cross-Linking in the Generation of a Marine-Mussel Adhesive. *Angew Chemie - Int Ed*. 2004;43(4):448-450. doi:10.1002/anie.200352759
39. Remko M, Fitz D, Rode BM. Effect of metal ions (Li+, Na+, K+, Mg2+, Ca2+, Ni2+, Cu2+, and Zn 2+) and water coordination on the structure and properties of L-arginine and zwitterionic L-arginine. *J Phys Chem A*. 2008;112(33):7652-7661. doi:10.1021/jp801418h
40. Hwang DS, Zeng H, Masic A, Harrington MJ, Israelachvili JN, Waite JH. Protein- and metal-dependent interactions of a prominent protein in mussel adhesive plaques. *J Biol Chem*. 2010;285(33):25850-25858. doi:10.1074/jbc.M110.133157
41. Wilker JJ. Marine bioinorganic materials: mussels pumping iron. *Curr Opin Chem Biol*. 2010;14(2):276-283. doi:10.1016/j.cbpa.2009.11.009
42. Hofman AH, van Hees IA, Yang J, Kamperman M. Bioinspired Underwater Adhesives by Using the Supramolecular Toolbox. *Adv Mater*. 2018;30(19). doi:10.1002/adma.201704640
43. Yu J, Wei W, Danner E, Ashley RK, Israelachvili JN, Waite JH. Mussel protein adhesion depends on interprotein thiol-mediated redox modulation. *Nat Chem Biol*. 2011;7(9):588-590. doi:10.1038/nchembio.630
44. Astoricchio E, Alfano C, Rajendran L, Temussi PA, Pastore A. The Wide World of Coacervates: From the Sea to Neurodegeneration. *Trends Biochem Sci*. 2020;45(8):706-717. doi:10.1016/j.tibs.2020.04.006
45. Rodriguez NRM, Das S, Kaufman Y, Israelachvili JN, Waite JH. Interfacial pH during mussel adhesive plaque formation. *Biofouling*. 2015;31(2):221-227. doi:10.1080/08927014.2015.1026337
46. Hwang DS, Sim SB, Cha HJ. Cell adhesion biomaterial based on mussel adhesive protein fused with RGD peptide. *Biomaterials*. 2007;28(28):4039-4046. doi:10.1016/j.biomaterials.2007.05.028

47. Jeon EY, Hwang BH, Yang YJ, et al. Rapidly light-activated surgical protein glue inspired by mussel adhesion and insect structural crosslinking. *Biomaterials*. 2015;67:11-19. doi:10.1016/j.biomaterials.2015.07.014
48. Hwang DS, Gim Y, Yoo HJ, Cha HJ. Practical recombinant hybrid mussel bioadhesive fp-151. *Biomaterials*. 2007;28(24):3560-3568. doi:10.1016/j.biomaterials.2007.04.039
49. Choi YS, Yang YJ, Yang B, Cha HJ. In vivo modification of tyrosine residues in recombinant mussel adhesive protein by tyrosinase co-expression in Escherichia coli. *Microb Cell Fact*. 2012;11:1-8. doi:10.1186/1475-2859-11-139
50. Guo Q, Chen J, Wang J, Zeng H, Yu J. Recent progress in synthesis and application of mussel-inspired adhesives. *Nanoscale*. 2020;12(3):1307-1324. doi:10.1039/c9nr09780e
51. Lee BP, Dalsin JL, Messersmith PB. Synthesis and gelation of DOPA-modified poly(ethylene glycol) hydrogels. *Biomacromolecules*. 2002;3(5):1038-1047. doi:10.1021/bm025546n
52. Zeng Z, Wang H, Morsi Y, Mo X. Synthesis and characterization of incorporating mussel mimetic moieties into photoactive hydrogel adhesive. *Colloids Surfaces B Biointerfaces*. 2018;161:94-102. doi:10.1016/j.colsurfb.2017.10.041
53. He Y, Sun C, Jiang F, et al. Lipids as integral components in mussel adhesion. *Soft Matter*. 2018;14(35):7145-7154. doi:10.1039/c8sm00509e
54. Petrone L, Kumar A, Sutanto CN, et al. Mussel adhesion is dictated by time-regulated secretion and molecular conformation of mussel adhesive proteins. *Nat Commun*. 2015;6:1-12. doi:10.1038/ncomms9737
55. Santonocito R, Venturella F, Piaz FD, et al. Recombinant mussel protein Pvfp-5β: A potential tissue bioadhesive. *J Biol Chem*. 2019;294(34):12826-12835. doi:10.1074/jbc.RA119.009531
56. Sreejamole KL, Radhakrishnan CK. Preliminary qualitative chemical evaluation of the extracts from mussel Perna viridis. *Int J Pharm Sci Rev Res*. 2010;5(2):38-42.
57. Jiang Z, Yu Y, Du L, et al. Peptide derived from Pvfp-1 as bioadhesive on bio-inert surface. *Colloids Surfaces B Biointerfaces*. 2012;90(1):227-235. doi:10.1016/j.colsurfb.2011.10.037

58. Liao Z, Jiang Y ting, Sun Q, Fan M hua, Wang J xin, Liang H ying. *Microstructure and In-Depth Proteomic Analysis of Perna Viridis Shell*. Vol 14.; 2019. doi:10.1371/journal.pone.0219699
59. Sainudin MS, Othman NH, Wan Mohammad WAS, Wan Ibrahim MH, Muthusamy K. Properties of concrete containing mussel (Perna viridis) shell ash as partial cement replacement. *Int J Integr Eng*. 2019;11(9 Special Issue):154-163.
60. Mag P, Krishnamoorthy V, Chuen LY, Sivayogi V, Kathiresan S, Bahari MB. Exploration of Antioxidant Capacity of Extracts of Perna viridis , a Marine Bivalve. Published online 2019. doi:10.4103/pm.pm
61. Yap C, Eugene Ng Y, Edward Thomas F, Cheng W, Ong G. The Use of Foot of the Green-Lipped Mussel is Perna Viridis as an Alternative Method to Reduce the Gender Effect on the Bioaccumulation of Cu and Zn in the Mussel. *Ann Limnol Oceanogr*. 2016;1(1):022-025. doi:10.17352/alo.000003
62. Ohkawa K, Nagai T, Nishida A, Yamomoto H. Purification of DOPA-containing foot proteins from green mussel, perna viridis, and adhesive properties of synthetic model copolypeptides. *J Adhes*. 2009;85(11):770-791. doi:10.1080/00218460903291353
63. Rajagopal S, Venugopalan VP, Van Der Velde G, Jenner HA. Greening of the coasts: A review of the Perna viridis success story. *Aquat Ecol*. 2006;40(3):273-297. doi:10.1007/s10452-006-9032-8
64. Hwang DS, Zeng H, Lu Q, Israelachvili J, Waite JH. Adhesion mechanism in a DOPA-deficient foot protein from green mussels. *Soft Matter*. 2012;8(20):5640-5648. doi:10.1039/c2sm25173f
65. Leung PTY, Ip JCH, Mak SST, et al. De novo transcriptome analysis of Perna viridis highlights tissue-specific patterns for environmental studies. *BMC Genomics*. 2014;15(1):1-14. doi:10.1186/1471-2164-15-804
66. Nishida A, Ohkawa K, Ueda I, Yamamoto H. Green mussel Perna viridis L.: Attachment behaviour and preparation of antifouling surfaces. *Biomol Eng*. 2003;20(4-6):381-387. doi:10.1016/S1389-0344(03)00057-1
67. Navinkumar J Patil, Paola Gabriela Vinueza Naranjo BZ. Wet Adhesive Properties of Asian Green Mussel (Perna viridis) Foot Protein Pvfp-5 : An Underwater Adhesive Primer Propiedades Adhesivas Húmedas de la Proteína del Pie de

Mejillones ... Wet Adhesive Properties of Asian Green Mussel (Perna viridis) Foot. 2018;1(April 2019):90-99.

68. Ou X, Xue B, Lao Y, et al. Structure and sequence features of mussel adhesive protein lead to its salt-tolerant adhesion ability. *Sci Adv*. 2020;6(39). doi:10.1126/sciadv.abb7620
69. Shin M, Shin JY, Kim K, et al. The position of lysine controls the catechol-mediated surface adhesion and cohesion in underwater mussel adhesion. *J Colloid Interface Sci*. 2020;563:168-176. doi:10.1016/j.jcis.2019.12.082
70. Bilotto P, Labate C, De Santo MP, Deepankumar K, Miserez A, Zappone B. Adhesive Properties of Adsorbed Layers of Two Recombinant Mussel Foot Proteins with Different Levels of DOPA and Tyrosine. *Langmuir*. 2019;35(48):15481-15490. doi:10.1021/acs.langmuir.9b01730
71. Overton TW. Recombinant protein production in bacterial hosts. *Drug Discov Today*. 2014;19(5):590-601. doi:10.1016/j.drudis.2013.11.008
72. Culotta VC, Wilkinson JK, Sollner-Webb B. Mouse and frog violate the paradigm of species-specific transcription of ribosomal RNA genes. *Proc Natl Acad Sci U S A*. 1987;84(21):7498-7502. doi:10.1073/pnas.84.21.7498
73. Quianzon CC, Cheikh I. History of insulin. *J Community Hosp Intern Med Perspect*. 2012;2(2):18701. doi:10.3402/jchimp.v2i2.18701
74. Koths K. Recombinant proteins for medical use: the attractions and challenges. *Curr Opin Biotechnol*. 1995;6(6):681-687. doi:10.1016/0958-1669(95)80112-X
75. Kadri K. Polymerase Chain Reaction (PCR): Principle and Applications. In: *Synthetic Biology - New Interdisciplinary Science*. IntechOpen; 2020. doi:10.5772/intechopen.86491
76. Harbers M. The current status of cDNA cloning. *Genomics*. 2008;91(3):232-242. doi:10.1016/j.ygeno.2007.11.004
77. Booth PJ, Curnow P. Folding scene investigation: membrane proteins. *Curr Opin Struct Biol*. 2009;19(1):8-13. doi:10.1016/j.sbi.2008.12.005
78. Block H, Maertens B, Spriestersbach A, et al. Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). A Review. In: *Methods in Enzymology*. Vol 463. Academic Press Inc.; 2009:439-473. doi:10.1016/S0076-6879(09)63027-5
79. Li MZ, Elledge SJ. SLIC: A method for sequence- and ligation-independent

cloning. *Methods Mol Biol*. 2012;852:51-59. doi:10.1007/978-1-61779-564-0_5

80. Hvidsten TR, Lægreid A, Kryshtafovych A, Anderson G, Fidelis K, Komorowski J. A comprehensive analysis of the structure-function relationship in proteins based on local structure similarity. *PLoS One*. 2009;4(7). doi:10.1371/journal.pone.0006266
81. Olshina MA, Sharon M. Mass spectrometry: A technique of many faces. *Q Rev Biophys*. 2016;49. doi:10.1017/S0033583516000160
82. Simmons TA, Green-Church KB, Limbach PA. Nucleic Acids and Nucleotides Studied Using Mass Spectrometry. In: *Encyclopedia of Spectroscopy and Spectrometry*. Elsevier; 1999:1681-1688. doi:10.1006/rwsp.2000.0213
83. Palmer T, Bonner PL. Monomeric and Oligomeric Enzymes. In: *Enzymes*. Elsevier; 2011:76-83. doi:10.1533/9780857099921.1.76
84. Hirst JD, Brooks CL. Helicity, circular dichroism and molecular dynamics of proteins. *J Mol Biol*. 1994;243(2):173-178. doi:10.1006/jmbi.1994.1644
85. Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochim Biophys Acta - Proteins Proteomics*. 2005;1751(2):119-139. doi:10.1016/j.bbapap.2005.06.005
86. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*. 2007;1(6):2876-2890. doi:10.1038/nprot.2006.202
87. Smyth MS, Martin JHJ. x Ray crystallography. *J Clin Pathol - Mol Pathol*. 2000;53(1):8-14. doi:10.1136/mp.53.1.8
88. Callaway E. "It opens up a whole new universe": Revolutionary microscopy technique sees individual atoms for first time. *Nature*. 2020;582(7811):156-157. doi:10.1038/d41586-020-01658-1
89. Juneja J, Udgaonkar JB. NMR studies of protein folding. *Curr Sci*. 2003;84(2):157-172.
90. Yang Y, Petkova A, Huang K, et al. An Achilles' heel in an amyloidogenic protein and its repair: Insulin fibrillation and therapeutic design. *J Biol Chem*. 2010;285(14):10806-10821. doi:10.1074/jbc.M109.067850
91. Liu G, Shen Y, Atreya HS, et al. NMR data collection and analysis protocol for high-throughput protein structure determination. *Proc Natl Acad Sci U S A*. 2005;102(30):10487-10492. doi:10.1073/pnas.0504338102

92. Neuhaus D. Nuclear Overhauser Effect. *Encycl Magn Reson*. Published online 2011:1-16. doi:10.1002/9780470034590.emrstm0350.pub2

93. Wu PSC, Ozawa K, Jergic S, Su XC, Dixon NE, Otting G. Amino-acid type identification in 15N-HSQC spectra by combinatorial selective 15N-labelling. *J Biomol NMR*. 2006;34(1):13-21. doi:10.1007/s10858-005-5021-9

94. Wittekind M, Mueller L. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *J Magn Reson Ser B*. 1993;101(2):201-205. doi:10.1006/jmrb.1993.1033

95. Yamazaki T, Muhandiranv DR, Kay LE, Lee W, Arrowsmith CH. A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. *J Am Chem Soc*. 1994;116(26):11655-11666. doi:10.1021/ja00105a005

96. Bax A, Ikura M. An efficient 3D NMR technique for correlating the proton and15N backbone amide resonances with the α-carbon of the preceding residue in uniformly15N/13C enriched proteins. *J Biomol NMR*. 1991;1(1):99-104. doi:10.1007/BF01874573

97. Brutscher B. Intraresidue HNCA and COHNCA experiments for protein backbone resonance assignment. *J Magn Reson*. 2002;156(1):155-159. doi:10.1006/jmre.2002.2546

98. Kay LE, Ikura M, Tschudin R, Bax A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson*. 1990;89(3):496-514. doi:10.1016/0022-2364(90)90333-5

99. Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J Biomol NMR*. 1993;3(2):185-204. doi:10.1007/BF00178261

100. Clore GM, Bax A, Driscoll PC, Gronenborn AM, Wingfield PT. Assignment of the Side-Chain 1H and 13C Resonances of Interleukin-1β Using Double-and Triple-Resonance Heteronuclear Three-Dimensional NMR Spectroscopy. *Biochemistry*. 1990;29(35):8172-8184. doi:10.1021/bi00487a027

101. King GF, Mobli M. Derivation of peptide and protein structure using NMR spectroscopy. In: *Comprehensive Natural Products II: Chemistry and Biology*. Vol

9. Elsevier Ltd; 2010:279-325. doi:10.1016/b978-008045382-8.00653-5

102. Baur M, Gemmecker G, Kessler H. 13C-NOESY-HSQC with Split Carbon Evolution for Increased Resolution with Uniformly Labeled Proteins. *J Magn Reson*. 1998;132(2):191-196. doi:10.1006/jmre.1998.1416
103. Yamazaki T, Forman-Kay JD, Kay LE. Two-Dimensional NMR Experiments for Correlating 13Cβ and 1Hδ/ϵ Chemical Shifts of Aromatic Residues in 13C-Labeled Proteins via Scalar Couplings. *J Am Chem Soc*. 1993;115(23):11054-11055. doi:10.1021/ja00076a099
104. Hu W, Kakalis LT, Jiang L, Jiang F, Ye X, Majumdar A. 3D HCCH-COSY-TOCSY experiment for the assignment of ribose and amino acid side chains in 13C labeled RNA and protein. *J Biomol NMR*. 1998;12(4):559-564. doi:10.1023/A:1008365301124
105. Mishra SH, Harden BJ, Frueh DP. A 3D time-shared NOESY experiment designed to provide optimal resolution for accurate assignment of NMR distance restraints in large proteins. *J Biomol NMR*. 2014;60(4):265-274. doi:10.1007/s10858-014-9873-8
106. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995;6(3):277-293. doi:10.1007/BF00197809
107. Vranken WF, Boucher W, Stevens TJ, et al. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct Funct Genet*. 2005;59(4):687-696. doi:10.1002/prot.20449
108. Rieping W, Bardiaux B, Bernard A, Malliavin TE, Nilges M. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics*. 2007;23(3):381-382. doi:10.1093/bioinformatics/btl589
109. Piotto M, Saudek V, Sklenář V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*. 1992;2(6):661-665. doi:10.1007/BF02192855
110. Mori S, Abeygunawardana C, Johnson MO, Van Zljl PCM. Improved Sensitivity of HSQC Spectra of Exchanging Protons at Short Interscan Delays Using a New Fast HSQC (FHSQC) Detection Scheme That Avoids Water Saturation. *J Magn Reson Ser B*. 1995;108(1):94-98. doi:10.1006/jmrb.1995.1109

111. Bodenhausen G, Ruben DJ. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett*. 1980;69(1):185-189. doi:10.1016/0009-2614(80)80041-8

112. Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ, Rance M. *Protein NMR Spectroscopy*. Elsevier Inc.; 2007. doi:10.1016/B978-0-12-164491-8.X5000-3

113. Grzesiek S, Bax A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J Magn Reson*. 1992;99(1):201-207. doi:10.1016/0022-2364(92)90169-8

114. Grzesiekt S, Bax A. Correlating Backbone Amide and Side Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. *J Am Chem Soc*. 1992;114(16):6291-6293. doi:10.1021/ja00042a003

115. Grzesiek S, Bax A. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson*. 1992;96(2):432-440. doi:10.1016/0022-2364(92)90099-S

116. Muhandiram DR, Kay LE. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *J Magn Reson Ser B*. 1994;103(3):203-216. doi:10.1006/jmrb.1994.1032

117. Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedures of uniformly $^{13}C/^{15}N$-enriched proteins. *J Biomol NMR*. 1993;3(2):185-204.

118. Lian LY, Roberts G. *Protein NMR Spectroscopy: Practical Techniques and Applications*. (Lian L-Y, Roberts G, eds.). John Wiley & Sons, Ltd; 2011. doi:10.1002/9781119972006

119. Gronenborn A. *'H-'H Correlation via Isotropic Mixing of 13C Magnetization, a New Three-Dimensional Approach for Assigning LH and 13C Spectra of 13C-Enriched Proteins*. Vol 88.; 1990.

120. Olejniczak ET, Xu RX, Fesik SW. A 4D HCCH-TOCSY experiment for assigning the side chain1H and13C resonances of proteins. *J Biomol NMR*. 1992;2(6):655-659. doi:10.1007/BF02192854

121. Gehring K, Ekiel I. H(C)CH-COSY and (H)CCH-COSY Experiments for 13C-Labeled Proteins in H2O Solution. *J Magn Reson*. 1998;135(1):185-193. doi:10.1006/jmre.1998.1543

122. Torizawa T, Ono AM, Terauchi T, Kainosho M. NMR assignment methods for the aromatic ring resonances of phenylalanine and tyrosine residues in proteins. *J Am Chem Soc*. 2005;127(36):12620-12626. doi:10.1021/ja051386m
123. Marion D, Driscoll PC, Kay LE, et al. Overcoming the Overlap Problem in the Assignment of 1H NMR Spectra of Larger Proteins by Use of Three-Dimensional Heteronuclear 1H-15N Hartmann-Hahn-Multiple Quantum Coherence and Nuclear Overhauser-Multiple Quantum Coherence Spectroscopy: Application to Interleukin 1β. *Biochemistry*. 1989;28(15):6150-6156. doi:10.1021/bi00441a004
124. Marion D, Kay LE, Sparks SW, Torchia DA, Bax A. Three-dimensional heteronuclear NMR of 15N-labeled proteins. *J Am Chem Soc*. 1989;111(4):1515-1517. doi:10.1021/ja00186a066
125. Zuiderweg ERP, Fesik SW. Heteronuclear Three-Dimensional NMR Spectroscopy of the Inflammatory Protein C5a. *Biochemistry*. 1989;28(6):2387-2391. doi:10.1021/bi00432a008
126. Chao FA, Byrd RA. Protein dynamics revealed by NMR relaxation methods. *Emerg Top Life Sci*. 2018;2(1):93-105. doi:10.1042/ETLS20170139
127. Lee D, Hilty C, Wider G, Wüthrich K. Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson*. 2006;178(1):72-76. doi:10.1016/j.jmr.2005.08.014
128. Gáspári Z, Perczel A. Protein Dynamics as Reported by NMR. *Annu Reports NMR Spectrosc*. 2010;71(C):35-75. doi:10.1016/B978-0-08-089054-8.00002-2
129. Baryshnikova OK, Sykes BD. Backbone dynamics of SDF-1α determined by NMR: Interpretation in the presence of monomer-dimer equilibrium. *Protein Sci*. 2006;15(11):2568-2578. doi:10.1110/ps.062255806
130. Kovermann M, Rogne P, Wolf-Watz M. Protein dynamics and function from solution state NMR spectroscopy. *Q Rev Biophys*. 2016;49:e6. doi:10.1017/S0033583516000019
131. Kay LE. NMR studies of protein structure and dynamics. *J Magn Reson*. 2005;173(2):193-207. doi:10.1016/j.jmr.2004.11.021
132. Mittermaier A, Kay LE. New tools provide new insights in NMR studies of protein dynamics. *Science (80-)*. 2006;312(5771):224-228. doi:10.1126/science.1124964
133. Bae S, Dyson HJ, Wright PE. Disordered Segments. *Mol Biol*. 2010;131(19):6814-

6821. doi:10.1021/ja809687r.Prediction

134. Jones R. 基因的改变NIH Public Access. *Bone*. 2014;23(1):1-7. doi:10.1016/j.bbapap.2010.10.012.An
135. Palmer AG. A dynamic look backward and forward. *J Magn Reson*. 2016;266:73-80. doi:10.1016/j.jmr.2016.01.018
136. *Dynamics Center User Manual Innovation with Integrity • NMR*. Accessed January 8, 2021. www.bruker.com
137. Barbato G, Ikura M, Kay LE, Bax A, Pastor RW. Backbone Dynamics of Calmodulin Studied by 15N Relaxation Using Inverse Detected Two-Dimensional NMR Spectroscopy: The Central Helix Is Flexible. *Biochemistry*. 1992;31(23):5269-5278. doi:10.1021/bi00138a005
138. Kay LE, Torchia DA, Bax A. Backbone Dynamics of Proteins As Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry*. 1989;28(23):8972-8979. doi:10.1021/bi00449a003
139. Kay LE, Nicholson LK, Delaglio F, Bax A, Torchia DA. Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins. *J Magn Reson*. 1992;97(2):359-375. doi:10.1016/0022-2364(92)90320-7
140. Palmer AG, Skelton NJ, Chazin WJ, Wright PE, Rance M. Suppression of the effects of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms in the measurement of spin-spin relaxation rates. *Mol Phys*. 1992;75(3):699-711. doi:10.1080/00268979200100511
141. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-3402. doi:10.1093/nar/25.17.3389
142. Weisshuhn PC, Sheppard D, Taylor P, et al. Non-Linear and Flexible Regions of the Human Notch1 Extracellular Domain Revealed by High-Resolution Structural Studies. *Structure*. 2016;24(4):555-566. doi:10.1016/j.str.2016.02.010
143. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: Protein structure and function prediction. *Nat Methods*. 2014;12(1):7-8.

doi:10.1038/nmeth.3213

144. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*. 2003;31(13):3381-3385. doi:10.1093/nar/gkg520
145. Barltrop JA, Owen TC, Cory AH, Cory JG. 5-(3-carboxymethoxyphenyl)-2-(4,5-dimethylthiazolyl)-3-(4-sulfophenyl)te trazolium, inner salt (MTS) and related analogs of 3-(4,5-dimethylthiazolyl)-2,5-diphenyltetrazolium bromide (MTT) reducing to purple water-soluble formazans As cell-viability indica. *Bioorganic Med Chem Lett*. 1991;1(11):611-614. doi:10.1016/S0960-894X(01)81162-8
146. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. *J Mol Biol*. 1963;7(1):95-99. doi:10.1016/S0022-2836(63)80023-6
147. Allain F, Mareuil F, Ménager H, Nilges M, Bardiaux B. ARIAweb: A server for automated NMR structure calculation. *Nucleic Acids Res*. 2020;48(1):W41-W47. doi:10.1093/nar/gkaa362
148. Mareuil F, Malliavin TE, Nilges M, Bardiaux B. Improved reliability, accuracy and quality in automated NMR structure calculation with ARIA. *J Biomol NMR*. 2015;62(4):425-438. doi:10.1007/s10858-015-9928-5
149. Bardiaux B, Malliavin T, Nilges M. *ARIA for Solution and Solid-State NMR*. Vol 831.; 2012. doi:10.1007/978-1-61779-480-3_23